GEOMETRIC APPROACH TO
CONVEX SUBDIFFERENTIAL CALCULUS

October 6, 2015

BORIS S. MORDUKHOVICH[1] and NGUYEN MAU NAM[2]

Dedicated to Franco Giannessi and Diethard Pallaschke with great respect

Abstract. In this paper we develop a geometric approach to convex subdifferential calculus in finite dimensions with employing some ideas of modern variational analysis. This approach allows us to obtain natural and rather easy proofs of basic results of convex subdifferential calculus in full generality and also derive new results of convex analysis concerning optimal value/marginal functions, normals to inverse images of sets under set-valued mappings, calculus rules for coderivatives of single-valued and set-valued mappings, and calculating coderivatives of solution maps to parameterized generalized equations governed by set-valued mappings with convex graphs.
Key words. convex analysis, generalized differentiation, geometric approach, convex separation, normal cone, subdifferential, coderivative, calculus rules, maximum function, optimal value function
AMS subject classifications. 49J52, 49J53, 90C31

1   Introduction

The notion of *subdifferential* (collection of *subgradients*) for nondifferentiable convex functions was independently introduced and developed by Moreau [15] and Rockafellar [19] who were both influenced by Fenchel [6]. Since then this notion has become one of the most central concepts of convex analysis and its various applications, including first of all convex optimization. The underlying difference between the standard derivative of a differentiable function and the subdifferential of a convex function at a given point is that the subdifferential is a *set* (of subgradients) which reduces to a singleton (gradient) if the function is differentiable. Due to the set-valuedness of the subdifferential, deriving *calculus rules* for it is a significantly more involved task in comparison with classical differential calculus. Needless to say that subdifferential calculus for convex functions is at the same level of importance as classical differential calculus, and it is difficult to imagine any usefulness of subgradients unless reasonable calculus rules are available.

The first and the most important result of convex subdifferential calculus is the subdifferential *sum rule*, which was obtained at the very beginning of convex analysis and has been since known as the *Moreau-Rockafellar theorem*. The reader can find this theorem and other results of convex subdifferential calculus in finite-dimensional spaces in the now classical monograph by Rockafellar [20]. More results in this direction in finite and infinite

dimensions with various applications to convex optimization, optimal control, numerical analysis, approximation theory, etc. are presented, e.g., in the monographs [1–3, 5, 7, 9–11, 14, 16, 17, 22] among the vast bibliography on the subject. In the recent time, convex analysis has become more and more important for applications to many new fields such as computational statistics, machine learning, and sparse optimization. Having this in mind, our major goal here is to revisit the convex subdifferential and provide an easy way to excess basic subdifferential calculus rules in finite dimensions.

In this paper, which can be considered as a supplement to our recent book [14], we develop a *a geometric approach* to convex subdifferential calculus. Our approach relies on the normal cone *intersection rule* based on *convex separation* and derives from it the major rules of subdifferential calculus without any appeal to duality, directional derivatives, and other tangentially generated constructions. This approach allows us to give direct and simple proofs of known results of convex subdifferential calculus in full generality and also to obtain some new results in this direction as those presented in Sections 9, 11, and 12. The developed approach is largely induced by the *dual-space* geometric approach of general variational analysis based on the *extremal principle* for systems of sets, which can be viewed as a variational counterpart of convex separation without necessarily imposing convexity; see [13] and the references therein.

Some of the results and proofs presented below have been outlined in the exercises of our book [14] while some other results (e.g., those related to subgradients of the optimal value function, coderivatives and their applications, etc.) seem to be new in the convex settings under consideration. In order to make the paper self-contained for the reader's convenience and also to make this material to be suitable for teaching, we recall here some basic definitions and properties of convex sets and functions with illustrative figures and examples. Our notation follows [14].

2 Basic Properties of Convex Sets

Here we recall some basic concepts and properties of convex sets. The detailed proofs of all the results presented in this and the next section can be found in [14]. Throughout the paper, consider the Euclidean space $\mathbb{R}^n$ of $n-$tuples of real numbers with the inner product

$$\langle x, y \rangle := \sum_{i=1}^{n} x_i y_i \ \text{ for } \ x = (x_1, \ldots, x_n) \in \mathbb{R}^n \text{ and } \ y = (y_1, \ldots, y_n) \in \mathbb{R}^n.$$

The Euclidean norm induced by this inner product is defined as usual by

$$\|x\| := \sqrt{\sum_{i=1}^{n} x_i^2}.$$

We often identify each element $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ with the column $x = [x_1, \ldots, x_n]^\top$.

Given two points $a, b \in \mathbb{R}^n$, the *line segment/interval* connecting a and b is

$$[a, b] := \{\lambda a + (1 - \lambda)b \mid \lambda \in [0, 1]\}.$$

A subset Ω of $\mathbb{R}^n$ is called *convex* if $\lambda x + (1 - \lambda)y \in \Omega$ for all $x, y \in \Omega$ and $\lambda \in (0, 1)$. A mapping $B\colon \mathbb{R}^n \to \mathbb{R}^p$ is *affine* if there exist a $p \times n$ matrix A and a vector $b \in \mathbb{R}^p$ such that

$$B(x) = Ax + b \text{ for all } x \in \mathbb{R}^n.$$

It is easy to check that the convexity of sets is preserved under images of affine mappings.

Proposition 2.1 *Let $B\colon \mathbb{R}^n \to \mathbb{R}^p$ be an affine mapping. The following properties hold:*

(i) *If Ω is a convex subset of $\mathbb{R}^n$, then the direct image $B(\Omega)$ is a convex subset of $\mathbb{R}^p$.*
(ii) *If Θ is a convex subset of $\mathbb{R}^p$, then the inverse image $B^{-1}(\Theta)$ is a convex subset of $\mathbb{R}^n$.*

For any collection of convex sets $\{\Omega_i\}_{i \in I}$, their intersection $\bigcap_{i \in I} \Omega_i$ is also convex. This motivates us to define the *convex hull* of a given set $\Omega \subset \mathbb{R}^n$ by

$$\mathrm{co}(\Omega) := \bigcap \Big\{ C \,\Big|\, C \text{ is convex and } \Omega \subset C \Big\},$$

i.e., the convex hull of a set Ω is the smallest convex set containing Ω. The following useful observation is a direct consequence of the definition.

Proposition 2.2 *For any subset Ω of $\mathbb{R}^n$, its convex hull admits the representation*

$$\mathrm{co}(\Omega) = \Big\{ \sum_{i=1}^{m} \lambda_i w_i \,\Big|\, \sum_{i=1}^{m} \lambda_i = 1,\ \lambda_i \geq 0,\ w_i \in \Omega,\ m \in \mathbb{N} \Big\},$$

where the symbol $\mathbb{N}$ stands for the set of all positive integers.

Given two points $a, b \in \mathbb{R}^n$, the *line* connecting a and b in $\mathbb{R}^n$ is defined by

$$L[a, b] := \{\lambda a + (1 - \lambda)b \mid \lambda \in \mathbb{R}\}.$$

A subset A of $\mathbb{R}^n$ is called *affine* if for any $x, y \in A$ and for any $\lambda \in \mathbb{R}$ we have

$$\lambda x + (1 - \lambda)y \in A,$$

which means that A is affine if and only if the line connecting any two points $a, b \in A$ is a subset of A. This shows that the intersection of any collection of affine sets is an affine set and thus allows us to define the *affine hull* of Ω by

$$\mathrm{aff}(\Omega) := \bigcap \Big\{ A \,\Big|\, A \text{ is affine and } \Omega \subset A \Big\}.$$

Similarly to the case of the convex hull, we have the following representation.

Proposition 2.3 *For any subset Ω of $\mathbb{R}^n$, its affine hull is represented by*

$$\mathrm{aff}(\Omega) = \Big\{ \sum_{i=1}^{m} \lambda_i \omega_i \,\Big|\, \sum_{i=1}^{m} \lambda_i = 1\ \omega_i \in \Omega,\ m \in \mathbb{N} \Big\}.$$
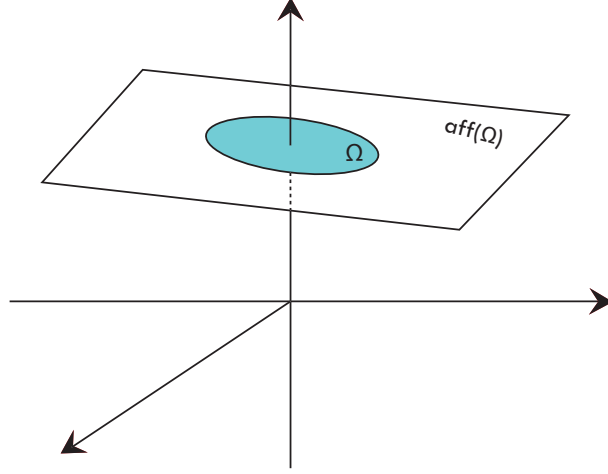
Figure 1: Affine hull.

Now we present some simple facts about affine sets.

Proposition 2.4 *Let A be an affine subset of $\mathbb{R}^n$. The following properties hold:*

(i) *If A contains 0, then it is a subspace of $\mathbb{R}^n$.*
(ii) *A is a closed, and so the affine hull of an arbitrary set Ω is always closed.*

Proof. **(i)** Since A is affine and since $0 \in A$, for any $x \in A$ and $\lambda \in \mathbb{R}$ we have that $\lambda x = \lambda x + (1 - \lambda)0 \in A$. It also holds

$$x + y = 2(x/2 + y/2) \in A$$

for any two elements $x, y \in A$, and thus A is a subspace of $\mathbb{R}^n$.

(ii) The conclusion is obvious if $A = \emptyset$. Suppose that $A \neq \emptyset$, choose $x_0 \in A$, and consider the set $L := A - x_0$. Then L is affine with $0 \in L$, and so it is a subspace of $\mathbb{R}^n$. Since $A = x_0 + L$ and any subspace of $\mathbb{R}^n$ is known to be closed, the set A is closed as well. $\square$

We are now ready to formulate the notion of the *relative interior* $\mathrm{ri}(\Omega)$ of a convex set $\Omega \subset \mathbb{R}^n$, which plays a central role in developing convex subdifferential calculus.

Definition 2.5 *We say that $x \in \mathrm{ri}(\Omega)$ if there exists $\gamma > 0$ such that*

$$\mathbb{B}(x; \gamma) \cap \mathrm{aff}(\Omega) \subset \Omega,$$

where $\mathbb{B}(x; \gamma)$ denotes the closed ball centered at x with radius γ.

The following simple proposition is useful in what follows and serves as an example for better understanding of the relative interior.
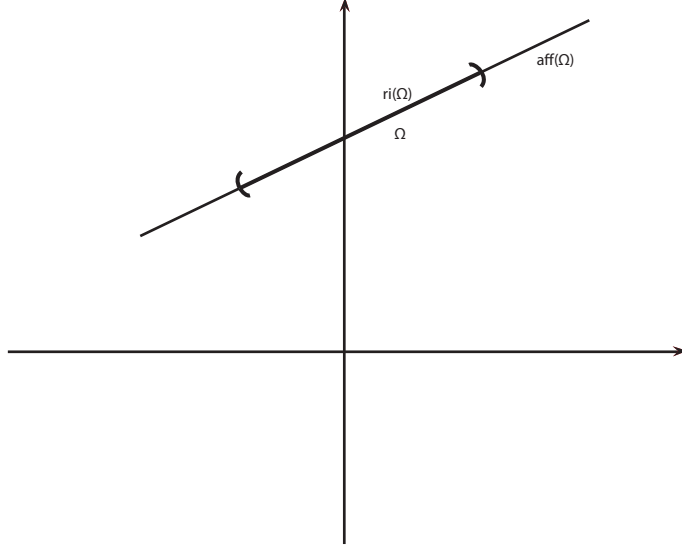
4

Figure 2: Relative interior.

Proposition 2.6 *Let Ω be a nonempty convex set. Suppose that $\bar{x} \in \text{ri}(\Omega)$ and $\bar{y} \in \Omega$. Then there exists $t > 0$ such that*

$$\bar{x} + t(\bar{x} - \bar{y}) \in \Omega.$$

Proof. Choose a number $\gamma > 0$ such that

$$\mathbb{B}(\bar{x}; \gamma) \cap \text{aff}(\Omega) \subset \Omega$$

and note that $\bar{x} + t(\bar{x} - \bar{y}) = (1 + t)\bar{x} + (-t)\bar{y} \in \text{aff}(\Omega)$ for all $t \in \mathbb{R}$ as it is an affine combination of $\bar{x}$ and $\bar{y}$. Select $t > 0$ so small that $\bar{x} + t(\bar{x} - \bar{y}) \in \mathbb{B}(\bar{x}; \gamma)$. Then we have $\bar{x} + t(\bar{x} - \bar{y}) \in \mathbb{B}(x; \gamma) \cap \text{aff}(\Omega) \subset \Omega$. $\qquad\qquad\square$

Given two elements $a, b \in \mathbb{R}^n$, define the half-open interval

$$[a, b) := \{ta + (1 - t)b \mid 0 < t \le 1\}.$$

The following theorem is taken from [20, Theorems 6.1 and 6.2]; see also [14, Theorem 1.72] for a direct and detailed proof.

Theorem 2.7 *Let Ω be a nonempty convex subset of $\mathbb{R}^n$. Then:*
(i) *We have $\text{ri}(\Omega) \ne \emptyset$ and*
(ii) *$[a, b) \subset \text{ri}(\Omega)$ for any $a \in \text{ri}(\Omega)$ and $b \in \overline{\Omega}$.*

The next theorem taken from [20, Theorem 6.7] gives us a convenient way to represent the relative interior of the direct image of a convex set under an affine mapping via of the relative interior of this set; see also [14, Excercise 1.27] and its solution for a simple proof.

5

Theorem 2.8 *Let $B : \mathbb{R}^n \to \mathbb{R}^p$ be affine, and let $\Omega \subset \mathbb{R}^n$ be convex. Then we have*

$$B(\mathrm{ri}(\Omega)) = \mathrm{ri}(B(\Omega)).$$

A useful consequence of this result is the following property concerning the *difference* of two subsets $A_1, A_2 \subset \mathbb{R}^n$ defined by

$$A_1 - A_2 := \{a_1 - a_2 \mid a_1 \in A_1 \text{ and } a_2 \in A_2\}.$$

Corollary 2.9 *Let Ω_1 and Ω_2 be convex subsets of $\mathbb{R}^n$. Then*

$$\mathrm{ri}(\Omega_1 - \Omega_2) = \mathrm{ri}(\Omega_1) - \mathrm{ri}(\Omega_2).$$

Proof. Consider the linear mapping $B \colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ given by $B(x, y) := x - y$ and form the Cartesian product $\Omega := \Omega_1 \times \Omega_2$. Then we have $B(\Omega) = \Omega_1 - \Omega_2$, which yields

$$\mathrm{ri}(\Omega_1 - \Omega_2) = \mathrm{ri}(B(\Omega)) = B(\mathrm{ri}(\Omega)) = B(\mathrm{ri}(\Omega_1 \times \Omega_2)) = B(\mathrm{ri}(\Omega_1) \times \mathrm{ri}(\Omega_2)) = \mathrm{ri}(\Omega_1) - \mathrm{ri}(\Omega_2)$$

by using the simple fact that $\mathrm{ri}(\Omega_1 \times \Omega_2) = \mathrm{ri}(\Omega_1) \times \mathrm{ri}(\Omega_2)$. $\qquad\square$

Given now a set $\Omega \subset \mathbb{R}^n$, the *distance function* associated with Ω is defined on $\mathbb{R}^n$ by

$$d(x; \Omega) := \inf\{\|x - \omega\| \mid \omega \in \Omega\}$$

and the *Euclidean projection* of x onto Ω is

$$\pi(x; \Omega) := \{\omega \in \Omega \mid \|x - \omega\| = d(x; \Omega)\}.$$

It is well known (see, e.g., [14, Corollary 1.76]) that $\pi(\bar{x}; \Omega)$ is a singleton whenever the set Ω is nonempty, closed, and convex in $\mathbb{R}^n$.

The next proposition plays a crucial role in proving major results on convex separation.

Proposition 2.10 *Let Ω be a nonempty closed convex subset of $\mathbb{R}^n$ with $\bar{x} \notin \Omega$. Then we have $\bar{\omega} = \pi(\bar{x}; \Omega)$ if and only if $\bar{\omega} \in \Omega$ and*

$$\langle \bar{x} - \bar{\omega}, \omega - \bar{\omega} \rangle \leq 0 \quad \text{for all } \omega \in \Omega. \tag{2.1}$$

Proof. Let us first show that (2.1) holds for $\bar{\omega} := \pi(\bar{x}; \Omega)$. Fixing any $t \in (0, 1)$ and $\omega \in \Omega$, we get $t\omega + (1 - t)\bar{\omega} \in \Omega$, which implies by the projection definition that

$$\|\bar{x} - \bar{\omega}\|^2 \leq \|\bar{x} - [t\omega + (1 - t)\bar{\omega}]\|^2.$$

This tells us by the construction of the Euclidean norm that

$$\|\bar{x} - \bar{\omega}\|^2 \leq \|\bar{x} - [\bar{\omega} + t(\omega - \bar{\omega})]\|^2 = \|\bar{x} - \bar{\omega}\|^2 - 2t\langle \bar{x} - \bar{\omega}, \omega - \bar{\omega} \rangle + t^2 \|\omega - \bar{\omega}\|^2$$

and yields therefore the inequality

$$2\langle \bar{x} - \bar{\omega}, \omega - \bar{\omega} \rangle \leq t\|\omega - \bar{\omega}\|^2.$$

Letting there $t \to 0^+$ justifies property (2.1).

Conversely, suppose that (2.1) is satisfied for $\bar{\omega} \in \Omega$ and get for any $\omega \in \Omega$ that

$$\|\bar{x} - \omega\|^2 = \|\bar{x} - \bar{\omega} + \bar{\omega} - \omega\|^2 = \|\bar{x} - \bar{\omega}\|^2 + 2\langle \bar{x} - \bar{\omega}, \bar{\omega} - \omega \rangle + \|\bar{\omega} - \omega\|^2 \geq \|\bar{x} - \bar{\omega}\|^2.$$

Thus we have $\|\bar{x} - \bar{\omega}\| \leq \|\bar{x} - \omega\|$ for all $\omega \in \Omega$, which verifies $\bar{\omega} = \pi(\bar{x}; \Omega)$. $\qquad\square$

3   Basic Properties of Convex Functions

In this section we deal with *extended-real-valued* functions $f\colon \mathbb{R}^n \to (-\infty, \infty] = \mathbb{R} \cup \{\infty\}$ and use the following arithmetic conventions on $(-\infty, \infty]$:

$$\alpha + \infty = \infty + \alpha = \infty \ \text{ for all } \ \alpha \in \mathbb{R},$$
$$\alpha \cdot \infty = \infty \cdot \alpha = \infty \text{ for all } \ \alpha > 0,$$
$$\infty + \infty = \infty, \quad \infty \cdot \infty = \infty, \quad 0 \cdot \infty = \infty \cdot 0 = 0.$$

The *domain* and *epigraph* of $f\colon \mathbb{R}^n \to (-\infty, \infty]$ are defined, respectively, by

$$\operatorname{dom}(f) := \{x \in \mathbb{R}^n \mid f(x) < \infty\}, \quad \operatorname{epi}(f) := \big\{(x, \alpha) \in \mathbb{R}^{n+1} \mid x \in \mathbb{R}^n, \ \alpha \geq f(x)\big\}.$$

Recall that a function $f\colon \mathbb{R}^n \to (-\infty, \infty]$ is *convex* on $\mathbb{R}^n$ if

$$f\big(\lambda x + (1 - \lambda)y\big) \leq \lambda f(x) + (1 - \lambda)f(y) \ \text{ for all } \ x, y \in \mathbb{R}^n \ \text{ and } \ \lambda \in (0, 1).$$

It is not hard to check that f is convex on $\mathbb{R}^n$ if and only if its epigraph is a convex subset of $\mathbb{R}^{n+1}$. Furthermore, the domain of a convex function is a convex set.

The class of convex functions is favorable for optimization theory and applications. The next proposition reveals a characteristic feature of convex functions from the viewpoint of minimization. Recall that f has a *local minimum* at $\bar{x} \in \operatorname{dom}(f)$ if there is $\gamma > 0$ such that

$$f(x) \geq f(\bar{x}) \text{ for all } \ x \in \mathbb{B}(\bar{x}; \gamma).$$

If this inequality holds for all $x \in \mathbb{R}^n$, we say that f has an *absolute/global minimum* at $\bar{x}$.

Proposition 3.1 *Let $f\colon \mathbb{R}^n \to (-\infty, \infty]$ be a convex function. Then f has a local minimum at $\bar{x}$ if and only if f has an absolute minimum at $\bar{x}$.*

Proof. We only need to show that any local minimizer of f provides a global minimum to this function. Suppose that $\bar{x}$ is such a local minimizer, fix any $x \in \mathbb{R}^n$, and denote $x_k := (1 - k^{-1})\bar{x} + k^{-1}x$ for all $k \in \mathbb{N}$. Then $x_k \to \bar{x}$ as $k \to \infty$. Taking $\gamma > 0$ from the definition of $\bar{x}$ gives us that $x_k \in \mathbb{B}(\bar{x}; \gamma)$ when k is sufficiently large. Hence

$$f(\bar{x}) \leq f(x_k) \leq (1 - k^{-1})f(\bar{x}) + k^{-1}f(x),$$

which readily implies that $f(\bar{x}) \leq f(x)$ whenever $x \in \mathbb{R}^n$. $\qquad \square$

Next we present the basic definition of the *subdifferential* as the collection of *subgradients* for a convex function at a given point.

Definition 3.2 *A vector $v \in \mathbb{R}^n$ is a subgradient of a convex function $f\colon \mathbb{R}^n \to (-\infty, \infty]$ at $\bar{x} \in \operatorname{dom}(f)$ if it satisfies the inequality*

$$f(x) \geq f(\bar{x}) + \langle v, x - \bar{x} \rangle \ \text{ for all } \ x \in \mathbb{R}^n.$$

The collection of subgradients is called the subdifferential of f at $\bar{x}$ and is denoted by $\partial f(\bar{x})$.
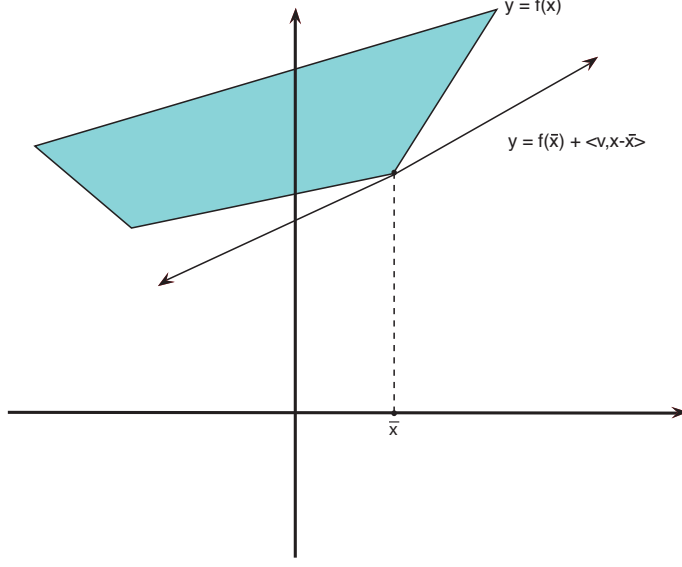
Figure 3: Subgradient.

Directly from the definition we have the *subdifferential Fermat rule*:

$$f \text{ has an absolute minimum at } \bar{x} \in \operatorname{dom}(f) \text{ if and only if } 0 \in \partial f(\bar{x}). \qquad (3.1)$$

Recall that $f \colon \mathbb{R}^n \to (-\infty, \infty]$ is (Fréchet) *differentiable* at $\bar{x} \in \operatorname{dom}(f)$ if there is a vector $v \in \mathbb{R}^n$ for which we have

$$\lim_{x \to \bar{x}} \frac{f(x) - f(\bar{x}) - \langle v, x - \bar{x} \rangle}{\|x - \bar{x}\|} = 0.$$

In this case the vector v is unique, is known as the *gradient* of f at $\bar{x}$, and is denoted by $\nabla f(\bar{x})$. The next proposition shows that the subdifferential of a convex function at a given point reduces to its gradient at this point when the function is differentiable.

Proposition 3.3 *Let $f \colon \mathbb{R}^n \to (-\infty, \infty]$ is convex and differentiable at $\bar{x} \in \operatorname{dom}(f)$. Then*

$$\langle \nabla f(\bar{x}), x - \bar{x} \rangle \leq f(x) - f(\bar{x}) \ \ \text{for all} \ \ x \in \mathbb{R}^n \ \ \text{with} \ \ \partial f(\bar{x}) = \{\nabla f(\bar{x})\}. \qquad (3.2)$$

Proof. Since f is differentiable at $\bar{x}$, for any $\epsilon > 0$ there exists $\gamma > 0$ such that

$$-\epsilon \|x - \bar{x}\| \leq f(x) - f(\bar{x}) - \langle \nabla f(\bar{x}), x - \bar{x} \rangle \leq \epsilon \|x - \bar{x}\| \text{ whenever } \|x - \bar{x}\| \leq \gamma.$$

Define further the function

$$\psi(x) := f(x) - f(\bar{x}) - \langle \nabla f(\bar{x}), x - \bar{x} \rangle + \epsilon \|x - \bar{x}\|$$

for which $\psi(x) \geq \psi(\bar{x}) = 0$ whenever $x \in \mathbb{B}(\bar{x}; \gamma)$. It follows from the convexity of ψ that $\psi(x) \geq \psi(\bar{x})$ when $x \in \mathbb{R}^n$, and thus

$$\langle \nabla f(\bar{x}), x - \bar{x} \rangle \leq f(x) - f(\bar{x}) + \epsilon \|x - \bar{x}\| \ \ \text{for all} \ \ x \in \mathbb{R}^n.$$

8

Letting now $\epsilon \to 0^+$ gives us the inequality in (3.2) and shows that $\nabla f(\bar{x}) \in \partial f(\bar{x})$.

To verify the remaining part of (3.2), pick any $v \in \partial f(\bar{x})$ and observe that

$$\langle v, x - \bar{x} \rangle \leq f(x) - f(\bar{x}) \text{ for all } x \in \mathbb{R}^n.$$

From the differentiability of f at $\bar{x}$ we have

$$\langle v - \nabla f(\bar{x}), x - \bar{x} \rangle \leq \epsilon \|x - \bar{x}\| \text{ whenever } \|x - \bar{x}\| \leq \gamma,$$

and so $\|v - \nabla f(\bar{x})\| \leq \epsilon$. This yields $v = \nabla f(\bar{x})$ since $\epsilon > 0$ is arbitrary and thus justifies the claimed relationship $\partial f(\bar{x}) = \{\nabla f(\bar{x})\}$. □

The following simple example calculates the subdifferential of the Euclidean norm function directly from the subdifferential definition.

Example 3.4 For the Euclidean norm function $p(x) := \|x\|$ we have

$$\partial p(x) = \begin{cases} \mathbb{B} & \text{if } x = 0, \\ \left\{ \dfrac{x}{\|x\|} \right\} & \text{otherwise,} \end{cases}$$

where $\mathbb{B}$ stands for the closed unit ball of $\mathbb{R}^n$. To verify this, we observe $\nabla p(x) = \dfrac{x}{\|x\|}$ for $x \neq 0$ due to the differentiability of $p(x)$ at nonzero points. Consider the case where $x = 0$ and use the definition describe $v \in \partial p(0)$ as

$$\langle v, x \rangle = \langle v, x - 0 \rangle \leq p(x) - p(0) = \|x\| \text{ for all } x \in \mathbb{R}^n.$$

For $x = v$ therein we get $\langle v, v \rangle \leq \|v\|$, which shows that $\|v\| \leq 1$, i.e., $v \in \mathbb{B}$. Conversely, picking $v \in \mathbb{B}$ and employing the Cauchy-Schwarz inequality tell us that

$$\langle v, x - 0 \rangle = \langle v, x \rangle \leq \|v\| \cdot \|x\| \leq \|x\| = p(x) - p(0) \text{ for all } x \in \mathbb{R}^n,$$

i.e., $v \in \partial p(0)$. Thus we arrive at the equality $\partial p(0) = \mathbb{B}$.

We conclude this section by the useful description of the relative interior of the graph of a convex function via the relative interior of its domain; see, e.g., [8, Proposition 1.1.9].

Proposition 3.5 *For a convex function* $f \colon \mathbb{R}^n \to (-\infty, \infty]$ *we have the representation*

$$\mathrm{ri}(\mathrm{epi}(f)) = \big\{ (x, \lambda) \mid x \in \mathrm{ri}(\mathrm{dom}(f)),\ \lambda > f(x) \big\}.$$

4 Convex Separation

Separation theorems for convex sets, which go back to Minkowski [12], have been well recognized among the most fundamental geometric tools of convex analysis. In this section we formulate and give simple proofs of several separation results needed in what follows under the weakest assumptions in finite dimensions. Let us begin with *strict separation* of a closed convex set and a point outside the set.

Proposition 4.1 *Let Ω be a nonempty closed convex set, and let $\bar{x} \notin \Omega$. Then there exists a nonzero vector $v \in \mathbb{R}^n$ such that*

$$\sup\{\langle v, x \rangle \mid x \in \Omega\} < \langle v, \bar{x} \rangle.$$

Proof. Denote $\bar{\omega} := \pi(\bar{x}; \Omega)$, $v := \bar{x} - \bar{\omega}$, and fix $x \in \Omega$. Proposition 2.10 gives us

$$\langle v, x - \bar{\omega} \rangle = \langle \bar{x} - \bar{\omega}, x - \bar{\omega} \rangle \leq 0,$$

which shows that

$$\langle v, x - \bar{\omega} \rangle = \langle v, x - \bar{x} + \bar{x} - \bar{\omega} \rangle = \langle v, x - \bar{x} + v \rangle \leq 0.$$

The last inequality therein yields

$$\langle v, x \rangle \leq \langle v, \bar{x} \rangle - \|v\|^2,$$

which implies in turn that

$$\sup\{\langle v, x \rangle \mid x \in \Omega\} < \langle v, \bar{x} \rangle$$

and thus completes the proof of the proposition. $\qquad\square$

Remark 4.2 It is easy to see that the closure $\overline{\Omega}$ of a convex set Ω is convex. If $\Omega \subset \mathbb{R}^n$ is a nonempty convex set with $\bar{x} \notin \overline{\Omega}$, then applying Proposition 4.1 to the convex set $\overline{\Omega}$ gives us a nonzero vector $v \in \mathbb{R}^n$ such that

$$\sup\{\langle v, x \rangle \mid x \in \Omega\} \leq \sup\{\langle v, x \rangle \mid x \in \overline{\Omega}\} < \langle v, \bar{x} \rangle.$$
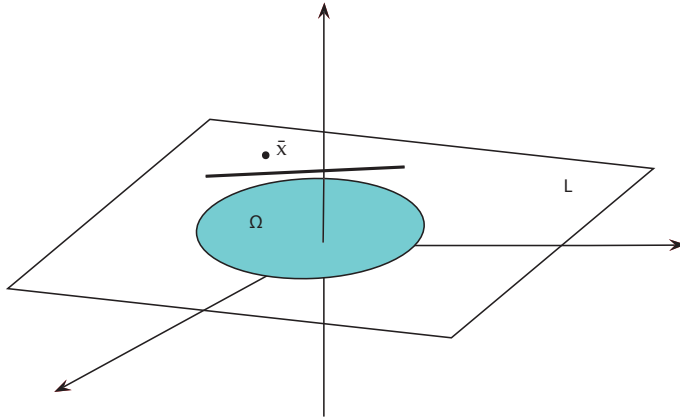

Figure 4: Separation in a subspace.

The next proposition justifies a strict separation property in a *subspace* of $\mathbb{R}^n$.

Proposition 4.3 *Let L be a subspace of $\mathbb{R}^n$, and let $\Omega \subset L$ be a nonempty convex set with $\bar{x} \in L$ and $\bar{x} \notin \overline{\Omega}$. Then there exists $v \in L$, $v \neq 0$, such that*

$$\sup\{\langle v, x \rangle \mid x \in \Omega\} < \langle v, \bar{x} \rangle.$$

Proof. Employing Remark 4.2 gives us a vector $w \in \mathbb{R}^n$ such that

$$\sup\{\langle w, x \rangle \mid x \in \Omega\} < \langle w, \bar{x} \rangle.$$

It is well known that $\mathbb{R}^n$ can be represented as the direct sum $\mathbb{R}^n = L \oplus L^{\perp}$, where

$$L^{\perp} := \{u \in \mathbb{R}^n \mid \langle u, x \rangle = 0 \ \text{ for all } \ x \in L\}.$$

Thus $w = u + v$ with $u \in L^{\perp}$ and $v \in L$. This yields $\langle u, x \rangle = 0$ for any $x \in \Omega \subset L$ and

$$\langle v, x \rangle = \langle u, x \rangle + \langle v, x \rangle = \langle u + v, x \rangle = \langle w, x \rangle \le \sup\{\langle w, x \rangle \mid x \in \Omega\}$$
$$< \langle w, \bar{x} \rangle = \langle u + v, \bar{x} \rangle = \langle u, \bar{x} \rangle + \langle v, \bar{x} \rangle = \langle v, \bar{x} \rangle,$$

which shows that $\sup\{\langle v, x \rangle \mid x \in \Omega\} < \langle v, \bar{x} \rangle$ with $v \ne 0$. $\qquad\qquad\square$
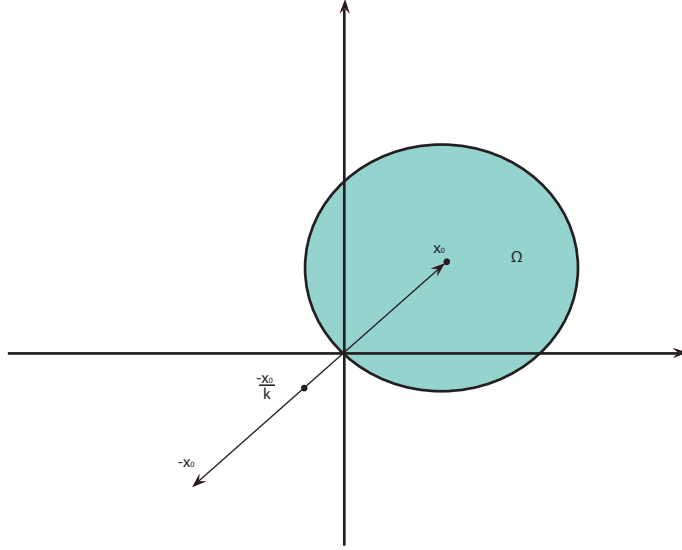

Figure 5: Illustration of the proof of Lemma 4.4.

Lemma 4.4 *Let $\Omega \subset \mathbb{R}^n$ be a nonempty convex set, and let $0 \in \overline{\Omega} \setminus \mathrm{ri}(\Omega)$. Then $\mathrm{aff}(\Omega)$ is a subspace of $\mathbb{R}^n$, and there is a sequence $\{x_k\} \subset \mathrm{aff}(\Omega)$ with $x_k \notin \overline{\Omega}$ and $x_k \to 0$ as $k \to \infty$.*

Proof. Since $\mathrm{ri}(\Omega) \ne \emptyset$ by Theorem 2.7(i) and $0 \in \overline{\Omega} \setminus \mathrm{ri}(\Omega)$, we find $x_0 \in \mathrm{ri}(\Omega)$ and conclude that $-tx_0 \notin \overline{\Omega}$ for all $t > 0$. Indeed, suppose by contradiction that $-tx_0 \in \overline{\Omega}$ for some $t > 0$ and then deduce from Theorem 2.7(ii) that

$$0 = \frac{t}{1+t}x_0 + \frac{1}{1+t}(-tx_0) \in \mathrm{ri}(\Omega),$$

which contradicts $0 \notin \mathrm{ri}(\Omega)$. Letting now $x_k := -\frac{x_0}{k}$ implies that $x_k \notin \overline{\Omega}$ for every k and $x_k \to 0$ as $k \to \infty$. Furthermore, we have

$$0 \in \overline{\Omega} \subset \overline{\mathrm{aff}(\Omega)} = \mathrm{aff}(\Omega)$$

11

by the closedness of $\mathrm{aff}(\Omega)$ due to Proposition 2.4(ii). This shows that $\mathrm{aff}(\Omega)$ is a subspace and that $x_k \in \mathrm{aff}(\Omega)$ for all $k \in \mathbb{N}$. $\qquad\square$

We continue with another important separation property known as *proper separation*.

Definition 4.5 *It is said that two nonempty convex sets Ω_1 and Ω_2 are properly separated if there exists a nonzero vector $v \in \mathbb{R}^n$ such that*

$$\sup\{\langle v, x\rangle \mid x \in \Omega_1\} \leq \inf\{\langle v, y\rangle \mid y \in \Omega_2\}, \quad \inf\{\langle v, x\rangle \mid x \in \Omega_1\} < \sup\{\langle v, y\rangle \mid y \in \Omega_2\}.$$

Lemma 4.6 *Let Ω be a nonempty convex set in $\mathbb{R}^n$. Then $0 \notin \mathrm{ri}(\Omega)$ if and only if the sets Ω and $\{0\}$ are properly separated, i.e., there is $v \in \mathbb{R}^n$, $v \neq 0$, such that*

$$\sup\{\langle v, x\rangle \mid x \in \Omega\} \leq 0, \quad \inf\{\langle v, x\rangle \mid x \in \Omega\} < 0.$$

Proof. We split the proof into the following two cases.

Case 1: $0 \notin \overline{\Omega}$. It follows from Remark 4.2 with $\bar{x} = 0$ that there exists $v \neq 0$ such that

$$\sup\{\langle v, x\rangle \mid x \in \Omega\} < \langle v, \bar{x}\rangle = 0,$$

and thus the sets Ω and $\{0\}$ are properly separated.

Case 2: $0 \in \overline{\Omega} \setminus \mathrm{ri}(\Omega)$. Letting $L := \mathrm{aff}(\Omega)$ and employing Lemma 4.4 tell us that L is a subspace of $\mathbb{R}^n$ and there is a sequence $\{x_k\} \subset L$ with $x_k \notin \overline{\Omega}$ and $x_k \to 0$ as $k \to \infty$. By Proposition 4.3 there is a sequence $\{v_k\} \subset L$ with $v_k \neq 0$ and

$$\sup\{\langle v_k, x\rangle \mid x \in \Omega\} < \langle v_k, x_k\rangle, \quad k \in \mathbb{N}.$$

Denoting $w_k := \frac{v_k}{\|v_k\|}$ shows that $\|w_k\| = 1$ for all $k \in \mathbb{N}$ and

$$\langle w_k, x\rangle < \langle w_k, x_k\rangle \text{ for all } x \in \Omega. \qquad (4.1)$$

Letting $k \to \infty$ in (4.1) and supposing without loss of generality that $w_k \to v \in L$ with some $\|v\| = 1$ along the whole sequence of $\{k\}$, we arrive at

$$\sup\{\langle v, x\rangle \mid x \in \Omega\} \leq 0$$

by taking into account that $|\langle w_k, x_k\rangle| \leq \|w_k\| \cdot \|x_k\| = \|x_k\| \to 0$. To verify further

$$\inf\{\langle v, x\rangle \mid x \in \Omega\} < 0,$$

it suffices to show that there is $x \in \Omega$ with $\langle v, x\rangle < 0$. Suppose by contradiction that $\langle v, x\rangle \geq 0$ for all $x \in \Omega$ and deduce from $\sup\{\langle v, x\rangle \mid x \in \Omega\} \leq 0$ that $\langle v, x\rangle = 0$ for all $x \in \Omega$. Since $v \in L = \mathrm{aff}(\Omega)$, we get the representation

$$v = \sum_{i=1}^{m} \lambda_i \omega_i \text{ with } \sum_{i=1}^{m} \lambda_i = 1 \text{ and } \omega_i \in \Omega \text{ for } i = 1, \dots, m,$$

which readily implies the equalities

$$\|v\|^2 = \langle v, v \rangle = \sum_{i=1}^{m} \lambda_i \langle v, \omega_i \rangle = 0$$

and so contradicts the condition $\|v\| = 1$. This justifies the proper separation of Ω and $\{0\}$.

To verify the reverse statement of the lemma, assume that Ω and $\{0\}$ are properly separated and thus find $0 \neq v \in \mathbb{R}^n$ such that

$$\sup\{\langle v, x \rangle \mid x \in \Omega\} \leq 0 \quad \text{while} \quad \langle v, \bar{x} \rangle < 0 \quad \text{for some} \quad \bar{x} \in \Omega.$$

Suppose by contradiction that $0 \in \operatorname{ri}(\Omega)$ and deduce from Proposition 2.6 that

$$0 + t(0 - \bar{x}) = -t\bar{x} \in \Omega \text{ for some } t > 0.$$

This immediately implies the inequalities

$$\langle v, -t\bar{x} \rangle \leq \sup\{\langle v, x \rangle \mid x \in \Omega\} \leq 0$$

showing that $\langle v, \bar{x} \rangle \geq 0$. It is a contradiction, which verifies $0 \notin \operatorname{ri}(\Omega)$. □

Now we are ready to prove the *main separation theorem* in convex analysis.

Theorem 4.7 *Let Ω_1 and Ω_2 be two nonempty convex subsets of $\mathbb{R}^n$. Then Ω_1 and Ω_2 are properly separated if and only if $\operatorname{ri}(\Omega_1) \cap \operatorname{ri}(\Omega_2) = \emptyset$.*

Proof. Define $\Omega := \Omega_1 - \Omega_2$ and verify that $\operatorname{ri}(\Omega_1) \cap \operatorname{ri}(\Omega_2) = \emptyset$ if and only if

$$0 \notin \operatorname{ri}(\Omega_1 - \Omega_2) = \operatorname{ri}(\Omega_1) - \operatorname{ri}(\Omega_2).$$

To proceed, suppose first that $\operatorname{ri}(\Omega_1) \cap \operatorname{ri}(\Omega_2) = \emptyset$ and so get by Corollary 2.9 that $0 \notin \operatorname{ri}(\Omega_1 - \Omega_2) = \operatorname{ri}(\Omega)$. Then Lemma 4.6 tells us that the sets Ω and $\{0\}$ are properly separated. Thus there exist $v \in \mathbb{R}^n$ with $\langle v, x \rangle \leq 0$ for all $x \in \Omega$ and also $y \in \Omega$ such that $\langle v, y \rangle < 0$. For any $\omega_1 \in \Omega_1$ and $\omega_2 \in \Omega_2$ we have $x := \omega_1 - \omega_2 \in \Omega$, and hence

$$\langle v, \omega_1 - \omega_2 \rangle = \langle v, x \rangle \leq 0,$$

which yields $\langle v, \omega_1 \rangle \leq \langle v, \omega_2 \rangle$. Choose $\bar{\omega}_1 \in \Omega_1$ and $\bar{\omega}_2 \in \Omega_2$ such that $y = \bar{\omega}_1 - \bar{\omega}_2$. Then

$$\langle v, \bar{\omega}_1 - \bar{\omega}_2 \rangle = \langle v, y \rangle < 0$$

telling us that $\langle v, \bar{\omega}_1 \rangle < \langle v, \bar{\omega}_2 \rangle$. Hence Ω_1 and Ω_2 are properly separated.

To prove next the converse implication, suppose that Ω_1 and Ω_2 are properly separated, which implies that the sets $\Omega = \Omega_1 - \Omega_2$ and $\{0\}$ are properly separated as well. Employing Lemma 4.6 again provides the relationships

$$0 \notin \operatorname{ri}(\Omega) = \operatorname{ri}(\Omega_1 - \Omega_2) = \operatorname{ri}(\Omega_1) - \operatorname{ri}(\Omega_2) \text{ and so } \operatorname{ri}(\Omega_1) \cap \operatorname{ri}(\Omega_2) = \emptyset,$$

which thus complete the proof of the theorem. □

Among various consequences of Theorem 4.7, including those presented below, note the following relationships between the closure and relative interior operations on convex sets that seemingly have nothing to do with separation.

13

Corollary 4.8 (i) *If $\Omega \subset \mathbb{R}^n$ is convex, then $\mathrm{ri}(\overline{\Omega}) = \mathrm{ri}(\Omega)$ and $\overline{\mathrm{ri}(\Omega)} = \overline{\Omega}$.*
(ii) *If both sets $\Omega_1, \Omega_2 \subset \mathbb{R}^n$ are convex, then we have $\mathrm{ri}(\Omega_1) = \mathrm{ri}(\Omega_2)$ provided that $\overline{\Omega}_1 = \overline{\Omega}_2$.*

Proof. (i) Both equalities are trivial if $\Omega = \emptyset$. To verify the first equality in (i) when $\Omega \neq \emptyset$, observe that for any $x \in \mathbb{R}^n$ we have the equivalence

$$\{x\} \text{ and } \Omega \text{ are properly separated } \iff \{x\} \text{ and } \overline{\Omega} \text{ are properly separated.}$$

Indeed, the implication "$\implies$" is obvious because for any $v \in \mathbb{R}^n$

$$\big[\sup\{\langle v, w\rangle \mid w \in \Omega\} \leq \langle v, x\rangle\big] \implies \big[\sup\{\langle v, w\rangle \mid w \in \overline{\Omega}\} \leq \langle v, x\rangle\big],$$

which can be proved by a limiting argument. For the converse, suppose that $\{x\}$ and $\overline{\Omega}$ are properly separated. Then there exists $v \in \mathbb{R}^n$, $v \neq 0$, such that

$$\sup\{\langle v, w\rangle \mid w \in \Omega\} \leq \langle v, x\rangle, \quad \inf\{\langle v, w\rangle \mid w \in \Omega\} < \langle v, x\rangle.$$

It follows that
$$\sup\{\langle v, w\rangle \mid w \in \Omega\} \leq \sup\{\langle v, w\rangle \mid w \in \overline{\Omega}\} \leq \langle v, x\rangle.$$

It remains to show that there exists $\bar{w} \in \Omega$ such that $\langle v, \bar{w}\rangle < \langle v, x\rangle$. If this is not the case, then $\langle v, w\rangle \geq \langle v, x\rangle$ for all $w \in \Omega$, which implies $\langle v, w\rangle = \langle v, x\rangle$ for all $w \in \Omega$. A simple limiting argument yields $\langle v, w\rangle = \langle v, x\rangle$ for all $w \in \overline{\Omega}$. This is a contradiction because $\inf\{\langle v, w\rangle \mid w \in \overline{\Omega}\} < \langle v, x\rangle$.

Defining further $\Theta := \{x\}$, we get $\mathrm{ri}(\Theta) = \{x\}$ and deduce from Theorem 4.7 that

$$\begin{aligned}
x \notin \mathrm{ri}(\overline{\Omega}) &\iff \mathrm{ri}(\Theta) \cap \mathrm{ri}(\overline{\Omega}) = \emptyset \\
&\iff \{x\} \text{ and } \overline{\Omega} \text{ are properly separated} \\
&\iff \{x\} \text{ and } \Omega \text{ are properly separated} \\
&\iff \mathrm{ri}(\Theta) \cap \mathrm{ri}(\Omega) = \emptyset \iff x \notin \mathrm{ri}(\Omega).
\end{aligned}$$

The second equality in (i) is a direct consequence of Theorem 2.7(ii).
(ii) If $\overline{\Omega}_1 = \overline{\Omega}_2$, then $\mathrm{ri}(\overline{\Omega}_1) = \mathrm{ri}(\overline{\Omega}_1)$ and hence $\mathrm{ri}(\Omega_1) = \mathrm{ri}(\Omega_2)$ by (i). $\qquad\square$

5   Normal Cone Intersection Rule

In this section we derive the central result of the geometric approach to convex subdifferential calculus, which provides a general *intersection rule* for the normal cone to convex sets. All the subsequent subdifferential results are consequences of this intersection rule.

Recall first the definition of the normal cone to a convex set.

Definition 5.1 *Let Ω be a nonempty convex subset of $\mathbb{R}^n$. Then the normal cone to the set Ω at $\bar{x} \in \Omega$ is defined by*

$$N(\bar{x}; \Omega) := \{v \in \mathbb{R}^n \mid \langle v, x - \bar{x}\rangle \leq 0 \ \text{ for all } \ x \in \Omega\}.$$

In the case where $\bar{x} \notin \Omega$ we define $N(\bar{x}; \Omega) := \emptyset$.
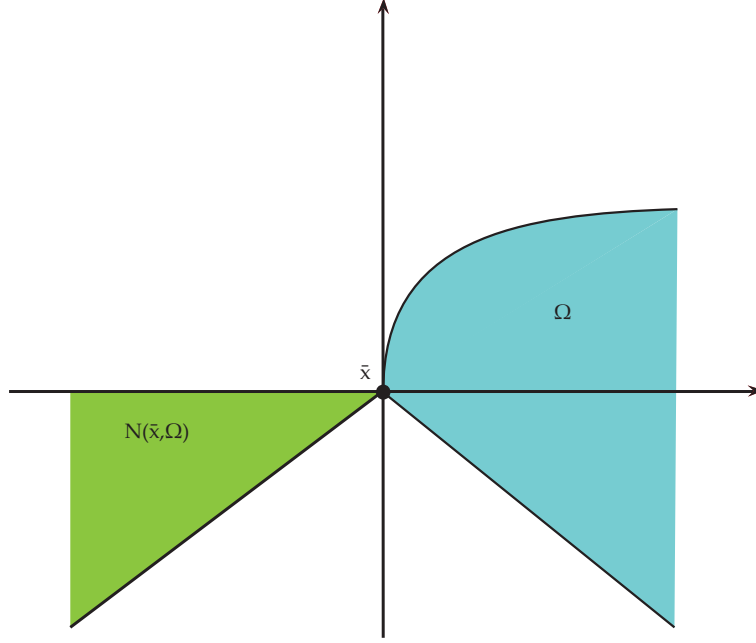
Figure 6: Normal cone.

It immediately follows from Definition 5.3 that $N(\bar{x}; \Omega)$ is a closed and convex cone, which reduces to $\{0\}$ if $\bar{x} \in \text{int}(\Omega)$. A remarkable property of the normal cone to a convex set Ω in finite dimensions is that $N(\bar{x}; \Omega) \neq \{0\}$ if and only if $\bar{x}$ is boundary point of Ω; see, e.g., [14, Corollary 2.14]. This is the classical *supporting hyperplane theorem*, which can be easily derived by the limiting procedure from Theorem 4.7.

Before deriving our major intersection result on the representation of the normal cone to finitely many convex sets, let us present a useful lemma on the relative interior of set intersections, which is also based on convex separation.

Lemma 5.2 *Let $\Omega_i \subset \mathbb{R}^n$ for $i = 1, \ldots, m$ with $m \geq 2$ be convex subsets of $\mathbb{R}^n$ such that*

$$\bigcap_{i=1}^{m} \text{ri}(\Omega_i) \neq \emptyset. \tag{5.1}$$

Then we have the representation

$$\text{ri}\Big(\bigcap_{i=1}^{m} \Omega_i\Big) = \bigcap_{i=1}^{m} \text{ri}(\Omega_i). \tag{5.2}$$

Proof. We first verify this result for $m = 2$. Pick $x \in \text{ri}(\Omega_1) \cap \text{ri}(\Omega_2)$ and find $\gamma > 0$ with

$$\mathbb{B}(x; \gamma) \cap \text{aff}(\Omega_1) \subset \Omega_1 \text{ and } \mathbb{B}(x; \gamma) \cap \text{aff}(\Omega_2) \subset \Omega_2,$$

which implies therefore that

$$\mathbb{B}(x; \gamma) \cap [\text{aff}(\Omega_1) \cap \text{aff}(\Omega_2)] \subset \Omega_1 \cap \Omega_2.$$

15

It is easy to see that $\mathrm{aff}(\Omega_1 \cap \Omega_2) \subset \mathrm{aff}(\Omega_1) \cap \mathrm{aff}(\Omega_2)$, and hence

$$\mathbb{B}(x; \gamma) \cap \mathrm{aff}(\Omega_1 \cap \Omega_2) \subset \Omega_1 \cap \Omega_2.$$

Thus we get $x \in \mathrm{ri}(\Omega_1 \cap \Omega_2)$, which justifies that $\mathrm{ri}(\Omega_1) \cap \mathrm{ri}(\Omega_2) \subset \mathrm{ri}(\Omega_1 \cap \Omega_2)$.

To verify the opposite inclusion in (5.2) for $m = 2$, observe that

$$\overline{\Omega_1 \cap \Omega_2} = \overline{\Omega}_1 \cap \overline{\Omega}_2 \tag{5.3}$$

for any convex sets Ω_1, Ω_2 with $\mathrm{ri}(\Omega_1) \cap \mathrm{ri}(\Omega_2) \neq \emptyset$. Indeed, pick $x \in \overline{\Omega}_1 \cap \overline{\Omega}_2$, $\bar{x} \in \mathrm{ri}(\Omega_1) \cap \mathrm{ri}(\Omega_2)$ and observe that $x_k := k^{-1}\bar{x} + (1 - k^{-1})x \to x$ as $k \to \infty$. Then Theorem 2.7(ii) tells us that $x_k \in \Omega_1 \cap \Omega_2$ for large $k \in \mathbb{N}$ and hence $x \in \overline{\Omega_1 \cap \Omega_2}$, which justifies the inclusion "$\supset$" in (5.3). The inclusion "$\subset$" therein obviously holds even for nonconvex sets. Now using (5.3) and the second equality in Corollary 4.8(i) gives us

$$\overline{\mathrm{ri}(\Omega_1) \cap \mathrm{ri}(\Omega_2)} = \overline{\mathrm{ri}(\Omega_1)} \cap \overline{\mathrm{ri}(\Omega_2)} = \overline{\Omega}_1 \cap \overline{\Omega}_2 = \overline{\Omega_1 \cap \Omega_2}.$$

Then we have the equality

$$\mathrm{ri}\left(\overline{\mathrm{ri}(\Omega_1) \cap \mathrm{ri}(\Omega_2)}\right) = \mathrm{ri}\left(\overline{\Omega_1 \cap \Omega_2}\right),$$

and thus conclude by Corollary 4.8(ii) that

$$\mathrm{ri}(\mathrm{ri}(\Omega_1) \cap \mathrm{ri}(\Omega_2)) = \mathrm{ri}(\Omega_1 \cap \Omega_2) \subset \mathrm{ri}(\Omega_1) \cap \mathrm{ri}(\Omega_2),$$

which justify representation (5.2) for $m = 2$.

To verify (5.2) under the validity of (5.1) in the general case of $m > 2$, we proceed by induction with taking into account that the result has been established for two sets and assuming that it holds for $m - 1$ sets. Considering m sets Ω_i, represent their intersection as

$$\bigcap_{i=1}^{m} \Omega_i = \Omega \cap \Omega_m \quad \text{with} \quad \Omega := \bigcap_{i=1}^{m-1} \Omega_i. \tag{5.4}$$

Then we have $\mathrm{ri}(\Omega) \cap \mathrm{ri}(\Omega_m) = \cap_{i=1}^{m} \mathrm{ri}(\Omega_i) \neq \emptyset$ by the imposed condition in (5.1) and the induction assumption on the validity of (5.2) for $m - 1$ sets. This allows us to employ the obtained result for the two sets Ω and Ω_m and thus arrive at the desired conclusion (5.2) for the m sets $\Omega_1, \ldots, \Omega_m$ under consideration. $\qquad \square$

Now we are ready to derive the underlying formula for the representation of the normal cone to intersections of finitely many convex sets. Note that the proof of this result and of the subsequent calculus rules for functions and set-valued mappings mainly follow the geometric pattern of variational analysis as in [13]. The specific features of convexity and the usage of convex separation instead of the extremal principle allow us to essentially simplify the proof and to *avoid the closedness* requirement on sets and the corresponding *lower semicontinuity* assumptions on functions in subdifferential calculus rules. Furthermore, we show below that the developed geometric approach works in the convex setting under the *relative interior qualification conditions*, which are well-recognized in finite-dimensional convex analysis and occur to be *weaker* than the basic/normal qualifications employed in [13, 14]; see. e.g., Corollary 5.5 and Example 5.6 below.
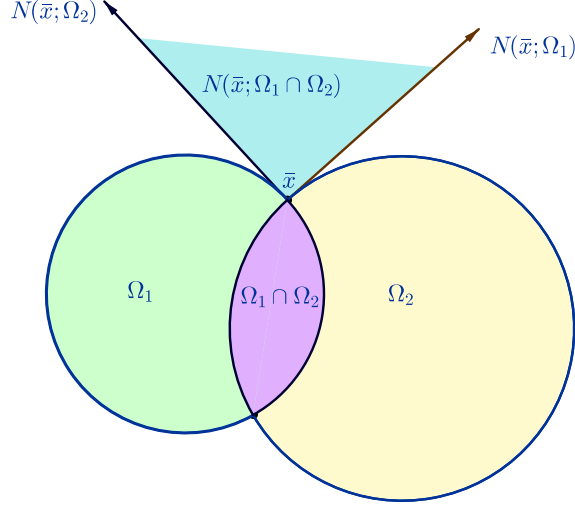
Figure 7: Intersection rule.

Theorem 5.3 *Let $\Omega_1, \ldots, \Omega_m \subset \mathbb{R}^n$ be convex sets satisfying the relative interior condition*

$$\bigcap_{i=1}^{m} \mathrm{ri}(\Omega_i) \neq \emptyset, \tag{5.5}$$

where $m \geq 2$. Then we have the intersection rule

$$N\Big(\bar{x}; \bigcap_{i=1}^{m} \Omega_i\Big) = \sum_{i=1}^{m} N(\bar{x}; \Omega_i) \text{ for all } \bar{x} \in \bigcap_{i=1}^{m} \Omega_i. \tag{5.6}$$

Proof. Proceeding by induction, let us first prove the statement of the theorem for the case of $m = 2$. Since the inclusion "$\supset$" in (5.6) trivially holds even without imposing (5.5), the real task is to verify the opposite inclusion therein. Fixing $\bar{x} \in \Omega_1 \cap \Omega_2$ and $v \in N(\bar{x}; \Omega_1 \cap \Omega_2)$, we get by the normal cone definition that

$$\langle v, x - \bar{x} \rangle \leq 0 \text{ for all } x \in \Omega_1 \cap \Omega_2.$$

Denote $\Theta_1 := \Omega_1 \times [0, \infty)$ and $\Theta_2 := \{(x, \lambda) \mid x \in \Omega_2, \ \lambda \leq \langle v, x - \bar{x} \rangle\}$. It follows from Proposition 3.5 that $\mathrm{ri}(\Theta_1) = \mathrm{ri}(\Omega_1) \times (0, \infty)$ and

$$\mathrm{ri}(\Theta_2) = \big\{(x, \lambda) \mid x \in \mathrm{ri}(\Omega_2), \ \lambda < \langle v, x - \bar{x} \rangle\big\}.$$

Arguing by contradiction, it is easy to check that $\mathrm{ri}(\Theta_1) \cap \mathrm{ri}(\Theta_2) = \emptyset$. Then applying Theorem 4.7 to these convex sets in $\mathbb{R}^{n+1}$ gives us $0 \neq (w, \gamma) \in \mathbb{R}^n \times \mathbb{R}$ such that

$$\langle w, x \rangle + \lambda_1 \gamma \leq \langle w, y \rangle + \lambda_2 \gamma \text{ for all } (x, \lambda_1) \in \Theta_1, \ (y, \lambda_2) \in \Theta_2. \tag{5.7}$$

Moreover, there are $(\tilde{x}, \tilde{\lambda}_1) \in \Theta_1$ and $(\tilde{y}, \tilde{\lambda}_2) \in \Theta_2$ satisfying

$$\langle w, \tilde{x} \rangle + \tilde{\lambda}_1 \gamma < \langle w, \tilde{y} \rangle + \tilde{\lambda}_2 \gamma.$$

17

Observe that $\gamma \le 0$ since otherwise we can get a contradiction by employing (5.7) with $(\bar{x}, k) \in \Theta_1$ as $k > 0$ and $(\bar{x}, 0) \in \Theta_2$. Let us now show by using (5.5) that $\gamma < 0$. Again arguing by contradiction, suppose that $\gamma = 0$ and then get

$$\langle w, x \rangle \le \langle w, y \rangle \ \text{ for all } \ x \in \Omega_1, \ y \in \Omega_2 \ \text{ and } \ \langle w, \tilde{x} \rangle < \langle w, \tilde{y} \rangle \ \text{ with } \ \tilde{x} \in \Omega_1, \ \tilde{y} \in \Omega_2.$$

This means the proper separation of the sets Ω_1 and Ω_2, which tells us by Theorem 4.7 that $\mathrm{ri}(\Omega_1) \cap \mathrm{ri}(\Omega_2) = \emptyset$. The obtained contradiction verifies the claim of $\gamma < 0$.

To proceed further, denote $\mu := -\gamma > 0$ and deduce from (5.7), by taking into account that $(x, 0) \in \Theta_1$ when $x \in \Omega_1$ and and that $(\bar{x}, 0) \in \Theta_2$, the inequality

$$\langle w, x \rangle \le \langle w, \bar{x} \rangle \ \text{ for all } \ x \in \Omega_1.$$

This yields $w \in N(\bar{x}; \Omega_1)$ and hence $\dfrac{w}{\mu} \in N(\bar{x}; \Omega_1)$. Moreover, we get from (5.7), due to $(\bar{x}, 0) \in \Theta_1$ and $(y, \alpha) \in \Theta_2$ for all $y \in \Omega_2$ with $\alpha = \langle v, y - \bar{x} \rangle$, that

$$\langle w, \bar{x} \rangle \le \langle w, y \rangle + \gamma \langle v, y - \bar{x} \rangle \ \text{ whenever } \ y \in \Omega_2.$$

Dividing both sides therein by γ, we arrive at the relationship

$$\left\langle \frac{w}{\gamma} + v, y - \bar{x} \right\rangle \le 0 \ \text{ for all } \ y \in \Omega_2,$$

and thus $\dfrac{w}{\gamma} + v = -\dfrac{w}{\mu} + v \in N(\bar{x}; \Omega_2)$. This gives us

$$v \in \frac{w}{\mu} + N(\bar{x}; \Omega_2) \subset N(\bar{x}; \Omega_1) + N(\bar{x}; \Omega_2)$$

completing therefore the proof of (5.6) in the case of $m = 2$.

Considering now the case of intersections for any finite number of sets, suppose by induction that the intersection rule (5.6) holds under (5.5) for $m - 1$ sets and verify that it continues to hold for the intersection of $m > 2$ sets $\bigcap_{i=1}^{m} \Omega_i$. Represent the latter intersection as $\Omega \cap \Omega_m$ with $\Omega := \bigcap_{i=1}^{m-1} \Omega_i$, we get from the imposed relative interior condition (5.5) and Lemma 5.2 that

$$\mathrm{ri}(\Omega) \cap \mathrm{ri}(\Omega_m) = \bigcap_{i=1}^{m} \mathrm{ri}(\Omega_i) \ne \emptyset.$$

Applying the intersection rule (5.6) to the two sets $\Omega \cap \Omega_m$ and then employing the induction assumption for $m - 1$ sets give us the equalities

$$N\left(\bar{x}; \bigcap_{i=1}^{m} \Omega_i\right) = N(\bar{x}; \Omega \cap \Omega_m) = N(\bar{x}; \Omega) + N(\bar{x}; \Omega_m) = \sum_{i=1}^{m} N(\bar{x}; \Omega_i),$$

which thus justify (5.6) for m sets and thus completes the proof of the theorem. $\qquad \square$

It is not difficult to observe the relative interior assumption (5.5) is *essential* for the validity of the intersection rule (5.6) as illustrated by the following example.

Example 5.4 Define the two convex sets on the plane by

$$\Omega_1 := \{(x, \lambda) \in \mathbb{R}^2 \mid \lambda \geq x^2\} \text{ and } \Omega_2 := \{(x, \lambda) \in \mathbb{R}^2 \mid \lambda \leq -x^2\}.$$

Then for $\bar{x} = (0, 0) \in \Omega_1 \cap \Omega_2$ we have

$$N(\bar{x}; \Omega_1) = \{0\} \times (-\infty, 0], \quad N(\bar{x}; \Omega_2) = \{0\} \times [0, \infty), \text{ and } N(\bar{x}; \Omega_1 \cap \Omega_2) = \mathbb{R}^2.$$

Thus $N(\bar{x}; \Omega_1) + N(\bar{x}; \Omega_2) = \{0\} \times \mathbb{R} \neq N(\bar{x}; \Omega_1 \cap \Omega_2)$, i.e., the intersection rule (5.6) fails. It does not contradict Theorem 5.3, since $\mathrm{ri}(\Omega_1) \cap \mathrm{ri}(\Omega_2) = \emptyset$ and so the relative interior qualification condition (5.5) does not hold in this case.
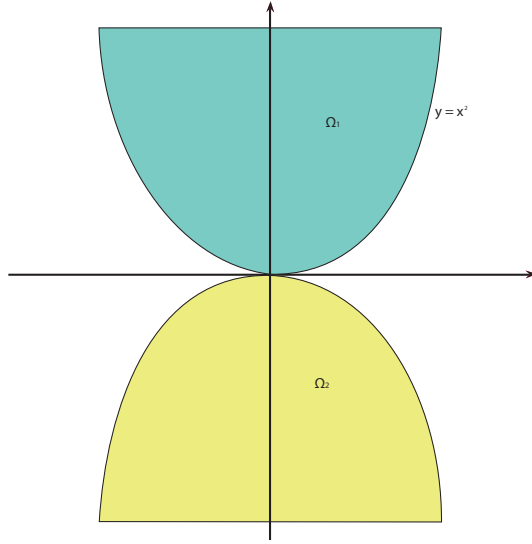

Figure 8: Illustration of the relative interior condition.

Finally, we compare the intersection rule of Theorem 5.3 derived under the relative interior qualification condition (5.5) with the corresponding result obtained in [14, Corollary 2.16] for $m = 2$ under the so-called *basic/normal qualification condition*

$$N(\bar{x}; \Omega_1) \cap [-N(\bar{x}; \Omega_2)] = \{0\} \tag{5.8}$$

introduced and applied earlier for deriving the intersection rule and related calculus results in nonconvex variational analysis; see, e.g., [13, 21] and the references therein. Let us first show that (5.8) yields (5.5) in the general convex setting.

Corollary 5.5 *Let $\Omega_1, \Omega_2 \subset \mathbb{R}^n$ be convex sets satisfying the basic qualification condition (5.8) at some $\bar{x} \in \Omega_1 \cap \Omega_2$. Then we have*

$$\mathrm{ri}(\Omega_1) \cap \mathrm{ri}(\Omega_2) \neq \emptyset, \tag{5.9}$$

and so the intersection rule (5.6) holds for these sets at any $\bar{x} \in \Omega_1 \cap \Omega_2$.

19

Proof. Arguing by contradiction, suppose that $\mathrm{ri}(\Omega_1) \cap \mathrm{ri}(\Omega_2) = \emptyset$. Then the sets Ω_1, Ω_2 are properly separated by Theorem 4.7, and so there is $v \neq 0$ such that

$$\langle v, x \rangle \leq \langle v, y \rangle \text{ for all } x \in \Omega_1, \ y \in \Omega_2.$$

Since $\bar{x} \in \Omega_2$, we have $\langle v, x - \bar{x} \rangle \leq 0$ for all $x \in \Omega_1$. Hence $v \in N(\bar{x}; \Omega_1)$ and similarly $-v \in N(\bar{x}; \Omega_2)$. Thus $0 \neq v \in N(\bar{x}; \Omega_1) \cap [-N(\bar{x}; \Omega_2)]$, which contradicts (5.8). $\qquad\square$

The next example demonstrates that (5.9) may be *strictly weaker* then (5.8).

Example 5.6 Consider the two convex sets on the plane defined by $\Omega_1 := \mathbb{R} \times \{0\}$ and $\Omega_2 := (-\infty, 0] \times \{0\}$. We obviously get that condition (5.9) is satisfied ensuring thus the validity of the intersection rule by Theorem 5.3. On the other hand, it follows for that

$$N(\bar{x}; \Omega_1) = \{0\} \times \mathbb{R} \text{ and } N(\bar{x}; \Omega_2) = [0, \infty) \times \mathbb{R} \text{ with } \bar{x} = (0, 0),$$

i.e., the other qualification condition (5.8) fails, which shows that the result of [14, Corollary 2.16] is not applicable in this case.

6 Subdifferential Sum Rule and Existence of Subgradients

The main goal of this section is to derive from the geometric intersection rule of Theorem 5.3 the subdifferential sum rule for convex extended-real-valued functions under the least restrictive relative interior qualification condition. Then we deduce from it a mild condition ensuring the existence of subgradients for general convex functions.

Prior to this, let us recall well-known relationships between normals to convex sets and subgradients of convex functions used in what follows.

Proposition 6.1 (i) *Let $\Omega \subset \mathbb{R}^n$ be a nonempty convex set, and let $\delta(x; \Omega) = \delta_\Omega(x)$ be its indicator function equal to 0 when $x \in \Omega$ and to ∞ otherwise. Then we have*

$$\partial \delta(\bar{x}; \Omega) = N(\bar{x}; \Omega) \ \text{ for any } \ \bar{x} \in \Omega.$$

(ii) *Let $f : \mathbb{R}^n \to (-\infty, \infty]$ be a convex function, and let $\bar{x} \in \mathrm{dom}(f)$. Then we have*

$$\partial f(\bar{x}) = \{v \in \mathbb{R}^n \mid (v, -1) \in N((\bar{x}, f(\bar{x})); \mathrm{epi}(f))\}.$$

Proof. (i) It follows directly from the definitions of the subdifferential, normal cone, and the set indicator function.

(ii) Fix any subgradient $v \in \partial f(\bar{x})$ and then get from Definition 3.2 that

$$\langle v, x - \bar{x} \rangle \leq f(x) - f(\bar{x}) \text{ for all } x \in \mathbb{R}^n. \tag{6.1}$$

To show that $(v, -1) \in N((\bar{x}, f(\bar{x})); \mathrm{epi}(f))$, fix any $(x, \lambda) \in \mathrm{epi}(f)$ and observe that due to $\lambda \geq f(x)$ we have the relationships

$$\langle (v, -1), (x, \lambda) - (\bar{x}, f(\bar{x})) \rangle = \langle v, x - \bar{x} \rangle + (-1)(\lambda - f(\bar{x}))$$
$$= \langle v, x - \bar{x} \rangle - (\lambda - f(\bar{x})) \leq \langle v, x - \bar{x} \rangle - (f(x) - f(\bar{x})) \leq 0,$$

where the the last inequality holds by (6.1). To verify the opposite inclusion in (ii), take $(v, -1) \in N((\bar{x}, f(\bar{x})); \mathrm{epi}(f))$ and fix any $x \in \mathrm{dom}(f)$. Then $(x, f(x)) \in \mathrm{epi}(f)$ and hence

$$\langle (v, -1), (x, f(x)) - (\bar{x}, f(\bar{x})) \rangle \leq 0,$$

which in turn implies the inequality

$$\langle v, x - \bar{x} \rangle - (f(x) - f(\bar{x})) \leq 0.$$

Thus $v \in \partial f(\bar{x})$, which completes the proof of the proposition. $\qquad\square$

Now we are ready to deduce the following subdifferential sum rule for function from the intersection rule of Theorem 5.3 for normals to sets.

Theorem 6.2 *Let $f_i \colon \mathbb{R}^n \to (-\infty, \infty]$, $i = 1, \ldots, m$, be extended-real-valued convex functions satisfying the relative interior qualification condition*

$$\bigcap_{i=1}^{m} \mathrm{ri}\Big(\mathrm{dom}(f_i)\Big) \neq \emptyset, \tag{6.2}$$

where $m \geq 2$. Then for all $\bar{x} \in \bigcap_{i=1}^{m} \mathrm{dom}(f_i)$ we have the sum rule

$$\partial\Big(\sum_{i=1}^{m} f_i\Big)(\bar{x}) = \sum_{i=1}^{m} \partial f_i(\bar{x}). \tag{6.3}$$

Proof. Observing that the inclusion "$\supset$" in (6.3) directly follows from the subdifferential definition, we proceed with the proof of the opposite inclusion. Consider first the case of $m = 2$ and pick any $v \in \partial(f_1 + f_2)(\bar{x})$. Then we have

$$\langle v, x - \bar{x} \rangle \leq (f_1 + f_2)(x) - (f_1 + f_2)(\bar{x}) \text{ for all } x \in \mathbb{R}^n. \tag{6.4}$$

Define the following convex subsets of $\mathbb{R}^{n+2}$ by

$$\Omega_1 := \{(x, \lambda_1, \lambda_2) \in \mathbb{R}^n \times \mathbb{R} \times \mathbb{R} \mid \lambda_1 \geq f_1(x)\} = \mathrm{epi}(f_1) \times \mathbb{R},$$
$$\Omega_2 := \{(x, \lambda_1, \lambda_2) \in \mathbb{R}^n \times \mathbb{R} \times \mathbb{R} \mid \lambda_2 \geq f_2(x)\}.$$

We can easily verify by (6.4) and the normal cone definition that

$$(v, -1, -1) \in N((\bar{x}, f_1(\bar{x}), f_2(\bar{x})); \Omega_1 \cap \Omega_2).$$

To apply Theorem 5.3 to these sets, let us check that $\mathrm{ri}(\Omega_1) \cap \mathrm{ri}(\Omega_2) \neq \emptyset$. Indeed, we get

$$\mathrm{ri}(\Omega_1) = \{(x, \lambda_1, \lambda_2) \in \mathbb{R}^n \times \mathbb{R} \times \mathbb{R} \mid x \in \mathrm{ri}(\mathrm{dom}(f_1)), \ \lambda_1 > f_1(x)\} = \mathrm{ri}(\mathrm{epi}(f_1) \times \mathbb{R},$$
$$\mathrm{ri}(\Omega_2) = \{(x, \lambda_1, \lambda_2) \in \mathbb{R}^n \times \mathbb{R} \times \mathbb{R} \mid x \in \mathrm{ri}(\mathrm{dom}(f_2)), \ \lambda_2 > f_2(x)\}$$

21

by Proposition 3.5. Then choosing $z \in \mathrm{ri}(\mathrm{dom}(f_1)) \cap \mathrm{ri}(\mathrm{dom}(f_2))$, it is not hard to see that

$$(z, f_1(z) + 1, f_2(z) + 1) \in \mathrm{ri}(\Omega_1) \cap \mathrm{ri}(\Omega_2) \neq \emptyset.$$

Applying now Theorem 5.3 to the above set intersection gives us

$$N((\bar{x}, f_1(\bar{x}), f_2(\bar{x})); \Omega_1 \cap \Omega_2) = N((\bar{x}, f_1(\bar{x}), f_2(\bar{x})); \Omega_1) + N((\bar{x}, f_1(\bar{x}), f_2(\bar{x})); \Omega_2).$$

It follows from the structures of the sets Ω_1 and Ω_2 that

$$(v, -1, -1) = (v_1, -\gamma_1, 0) + (v_2, 0, -\gamma_2)$$

with $(v_1, -\gamma_1) \in N((\bar{x}, f_1(\bar{x})); \mathrm{epi}(f_1))$ and $(v_2, -\gamma_2) \in N((\bar{x}, f_2(\bar{x})); \mathrm{epi}(f_2))$. Thus

$$v = v_1 + v_2, \quad \gamma_1 = \gamma_2 = 1,$$

and we have by Proposition 6.1(ii) that $v_1 \in \partial f_1(\bar{x})$ and $v_2 \in \partial f_2(\bar{x})$. This ensures the inclusion $\partial(f_1 + f_2)(\bar{x}) \subset \partial f_1(\bar{x}) + \partial f_2(\bar{x})$ and hence verifies (6.3) in the case of $m = 2$. To complete the proof of the theorem in the general case of $m > 2$, we proceed by induction similarly to the proof of Theorem 5.3 with the usage of Lemma 5.2 to deal with relative interiors in the qualification condition (6.2). □

The next result is a simple consequence of Theorem 6.2 providing a mild condition for the existence of subgradients of an extended-real-valued convex function at a given point.

Corollary 6.3 *Let $f \colon \mathbb{R}^n \to (-\infty, \infty]$ be a convex function. Then the validity of the relative interiority condition $\bar{x} \in \mathrm{ri}(\mathrm{dom}\, f)$ ensures that $\partial f(\bar{x}) \neq \emptyset$.*

Proof. Define the extended-real-valued function on $\mathbb{R}^n$ by

$$g(x) := f(x) + \delta_{\{\bar{x}\}}(x) = \begin{cases} f(\bar{x}) & \text{if } x = \bar{x}, \\ \infty & \text{otherwise} \end{cases}$$

via the indicator function of the singleton $\{\bar{x}\}$. Then $\mathrm{epi}(g) = \{\bar{x}\} \times [f(\bar{x}), \infty)$ and hence $N((\bar{x}, g(\bar{x})); \mathrm{epi}(g)) = \mathbb{R}^n \times (-\infty, 0]$. We obviously get that $\partial g(\bar{x}) = \mathbb{R}^n$ and that $\partial \delta_{\{\bar{x}\}}(\bar{x}) = N(\bar{x}; \{\bar{x}\}) = \mathbb{R}^n$ by Proposition 6.1(i). We further have

$$\mathrm{ri}(\mathrm{dom}(h)) = \{\bar{x}\} \quad \text{for} \quad h(x) := \delta_{\{\bar{x}\}}(x)$$

and thus $\mathrm{ri}(\mathrm{dom}(f)) \cap \mathrm{ri}(\mathrm{dom}(h)) \neq \emptyset$. Applying the subdifferential sum rule of Theorem 6.2 to the above function $g(x)$ at $\bar{x}$ gives us

$$\mathbb{R}^n = \partial g(\bar{x}) = \partial f(\bar{x}) + \mathbb{R}^n,$$

which justifies the claimed assertion on $\partial f(\bar{x}) \neq \emptyset$. □

7 Subdifferential Chain Rule

In this section we employ the intersection rule of Theorem 5.3 to derive a chain rule of the subdifferential of a composition of an extended-real-valued function and an affine mapping under which we obviously keep convexity. First we present the following useful lemma.

Lemma 7.1 *Let $B : \mathbb{R}^n \to \mathbb{R}^p$ be an affine mapping given by $B(x) := Ax + b$, where A is a $p \times n$ matrix and $b \in \mathbb{R}^p$. Then for any $(\bar{x}, \bar{y}) \in \mathrm{gph}(B)$ we have*

$$N\big((\bar{x}, \bar{y}); \mathrm{gph}(B)\big) = \big\{(u, v) \in \mathbb{R}^n \times \mathbb{R}^p \mid u = -A^\top v\big\}.$$

Proof. It is clear that $\mathrm{gph}(B)$ is convex and $(u, v) \in N((\bar{x}, \bar{y})\,\mathrm{gph}(B))$ if and only if

$$\langle u, x - \bar{x} \rangle + \langle v, B(x) - B(\bar{x}) \rangle \le 0 \ \text{ for all } \ x \in \mathbb{R}^n. \tag{7.1}$$

It follows directly from the definitions that

$$\begin{aligned}
\langle u, x - \bar{x} \rangle + \langle v, B(x) - B(\bar{x}) \rangle &= \langle u, x - \bar{x} \rangle + \langle v, A(x) - A(\bar{x}) \rangle \\
&= \langle u, x - \bar{x} \rangle + \langle A^\top v, x - \bar{x} \rangle = \langle u + A^\top v, x - \bar{x} \rangle.
\end{aligned}$$

This implies the equivalence of (7.1) to $\langle u + A^\top v, x - \bar{x} \rangle \le 0$ for all $x \in \mathbb{R}^n$, and so to $u = -A^\top v$. $\qquad\square$

Theorem 7.2 *Let $f : \mathbb{R}^p \to (-\infty, \infty]$ be a convex function, and let $B : \mathbb{R}^n \to \mathbb{R}^p$ be as in Lemma 7.1 with $B(\bar{x}) \in \mathrm{dom}(f)$ for some $\bar{x} \in \mathbb{R}^n$. Denote $\bar{y} := B(\bar{x})$ and assume that the range of B contains a point of $\mathrm{ri}(\mathrm{dom}(f))$. Then we have the subdifferential chain rule*

$$\partial (f \circ B)(\bar{x}) = A^\top \big(\partial f(\bar{y})\big) = \big\{ A^\top v \mid v \in \partial f(\bar{y}) \big\}. \tag{7.2}$$

Proof. Fix $v \in \partial (f \circ B)(\bar{x})$ and form the subsets of $\mathbb{R}^n \times \mathbb{R}^p \times \mathbb{R}$ by

$$\Omega_1 := \mathrm{gph}(B) \times \mathbb{R} \ \text{ and } \ \Omega_2 := \mathbb{R}^n \times \mathrm{epi}(f).$$

Then we clearly get the relationships

$$\mathrm{ri}(\Omega_1) = \Omega_1 = \mathrm{gph}(B) \times \mathbb{R}, \quad \mathrm{ri}(\Omega_2) = \{(x, y, \lambda) \mid x \in \mathbb{R}^n, \ y \in \mathrm{ri}(\mathrm{dom}(f)), \ \lambda > f(y)\},$$

and thus the assumption of the theorem tells us that $\mathrm{ri}(\Omega_1) \cap \mathrm{ri}(\Omega_2) \ne \emptyset$.

Further, it follows from the definitions of the subdifferential and of the normal cone that $(v, 0, -1) \in N((\bar{x}, \bar{y}, \bar{z}); \Omega_1 \cap \Omega_2)$, where $\bar{z} := f(\bar{y})$. Indeed, for any $(x, y, \lambda) \in \Omega_1 \cap \Omega_2$ we have $y = B(x)$ and $\lambda \ge f(y)$, and so $\lambda \ge f(B(x))$. Thus

$$\langle v, x - \bar{x} \rangle + 0(y - \bar{y}) + (-1)(\lambda - \bar{z}) \le \langle v, x - \bar{x} \rangle - [f(B(x)) - f(B(\bar{x}))] \le 0.$$

Employing the intersection rule of Theorem 5.3 to the above sets gives us

$$(v, 0, -1) \in N\big((\bar{x}, \bar{y}, \bar{z}); \Omega_1\big) + N\big((\bar{x}, \bar{y}, \bar{z}); \Omega_2\big),$$

23

which reads that $(v, 0, -1) = (v, -w, 0) + (0, w, -1)$ with $(v, -w) \in N((\bar{x}, \bar{y}); \mathrm{gph}(B))$ and $(w, -1) \in N((\bar{y}, \bar{z}); \mathrm{epi}(f))$. Then we get

$$v = A^\top w \quad \text{and} \quad w \in \partial f(\bar{y}),$$

which implies in turn that $v \in A^\top(\partial f(\bar{y}))$ and hence verifies the inclusion "⊂" in (7.2). The opposite inclusion follows directly from the definition of the subdifferential. □

8   Subdifferentiation of Maximum Functions

Our next topic is subdifferentiation of an important class of nonsmooth convex functions defined as the pointwise maximum of convex functions. We calculate the subdifferential of such functions by using again the geometric intersection rule of Theorem 5.3.

Given $f_i \colon \mathbb{R}^n \to (-\infty, \infty]$ for $i = 1, \ldots, m$, define the *maximum function* by

$$f(x) := \max_{i=1,\ldots,m} f_i(x), \quad x \in \mathbb{R}^n, \tag{8.1}$$

and for $\bar{x} \in \mathbb{R}^n$ consider the *active index* set

$$I(\bar{x}) := \{ i \in \{1, \ldots, m\} \mid f_i(\bar{x}) = f(\bar{x}) \}.$$

Lemma 8.1 (i) *Let Ω be a convex set in $\mathbb{R}^n$. Then $\mathrm{int}\,(\Omega) = \mathrm{ri}(\Omega)$ provided that $\mathrm{int}\,(\Omega) \neq \emptyset$. Furthermore, $N(\bar{x}; \Omega) = \{0\}$ if $\bar{x} \in \mathrm{int}(\Omega)$.*
(ii) *Let $f \colon \mathbb{R}^n \to (-\infty, \infty]$ be a convex function, which is continuous at $\bar{x} \in \mathrm{dom}(f)$. Then we have $\bar{x} \in \mathrm{int}(\mathrm{dom}(f))$ with the implication*

$$(v, -\lambda) \in N((\bar{x}, f(\bar{x})); \mathrm{epi}(f)) \implies [\lambda \geq 0 \quad \text{and} \quad v \in \lambda \partial f(\bar{x})].$$

Proof. (i) Suppose that $\mathrm{int}(\Omega) \neq \emptyset$ and check that $\mathrm{aff}(\Omega) = \mathbb{R}^n$. Indeed, picking $\bar{x} \in \mathrm{int}(\Omega)$ and fixing $x \in \mathbb{R}^n$, find $t > 0$ with $tx + (1-t)\bar{x} = \bar{x} + t(x - \bar{x}) \in \mathrm{int}(\Omega) \subset \mathrm{aff}(\Omega)$. It yields

$$x = \frac{1}{t}(tx + (1-t)\bar{x}) + (1 - \frac{1}{t})\bar{x} \in \mathrm{aff}(\Omega),$$

which justifies the claimed statement due to the definition of relative interior.

To verify the second statement in (i), take $v \in N(\bar{x}; \Omega)$ with $\bar{x} \in \mathrm{int}(\Omega)$ and get

$$\langle v, x - \bar{x} \rangle \leq 0 \text{ for all } x \in \Omega.$$

Choosing $\delta > 0$ such that $\bar{x} + tv \in \mathbb{B}(\bar{x}; \delta) \subset \Omega$ for $t > 0$ sufficiently small, gives us

$$\langle v, \bar{x} + tv - \bar{x} \rangle = t\|v\|^2 \leq 0,$$

which implies $v = 0$ and thus completes the proof of assertion (i).
(ii) The continuity of f allows us to find $\delta > 0$ such that

$$|f(x) - f(\bar{x})| < 1 \text{ whenever } x \in \mathbb{B}(\bar{x}; \delta).$$

This yields $\mathbb{B}(\bar{x}; \delta) \subset \text{dom}(f)$ and shows therefore that $\bar{x} \in \text{int}(\text{dom}(f))$.

Now suppose that $(v, -\lambda) \in N((\bar{x}, f(\bar{x})); \text{epi}(f))$. Then

$$\langle v, x - \bar{x} \rangle - \lambda(t - f(\bar{x})) \leq 0 \text{ whenever } (x, t) \in \text{epi}(f). \tag{8.2}$$

Employing this inequality with $x = \bar{x}$ and $t = f(\bar{x}) + 1$ yields $\lambda \geq 0$.

If $\lambda > 0$, we readily get $(v/\lambda, -1) \in N((\bar{x}, f(\bar{x})); \text{epi}(f))$. It follows from Proposition 6.1 that $v/\lambda \in \partial f(\bar{x})$, and hence $v \in \lambda \partial f(\bar{x})$.

In the case where $\lambda = 0$, we deduce from (8.2) that $v \in N(\bar{x}; (\text{dom}(f)) = \{0\}$, and so the inclusion $v \in \lambda \partial f(\bar{x})$ is also valid. Note that $\partial f(\bar{x}) \neq \emptyset$ by Corollary 6.3. $\qquad \square$

Now we are ready to derive the following *maximum rule*.

Theorem 8.2 *Let* $f_i \colon \mathbb{R}^n \to (-\infty, \infty]$, $i = 1, \ldots, m$, *be convex functions, and let* $\bar{x} \in \bigcap_{i=1}^m \text{dom} f_i$ *be such that each* f_i *is continuous at* $\bar{x}$. *Then we have the maximum rule:*

$$\partial\big(\max f_i\big)(\bar{x}) = \text{co}\left(\bigcup_{i \in I(\bar{x})} \partial f_i(\bar{x})\right).$$

Proof. Let f be the maximum function defined in (8.1) for which we obviously have

$$\text{epi}(f) = \bigcap_{i=1}^m \text{epi}(f_i).$$

Employing Proposition 3.5 and Lemma 8.1(i) give us the equalities

$$\text{ri}(\text{epi}(f_i)) = \{(x, \lambda) \mid x \in \text{ri}(\text{dom}(f_i)), \ \lambda > f_i(x)\} = \{(x, \lambda) \mid x \in \text{int}(\text{dom}(f_i)), \ \lambda > f_i(x)\},$$

which imply that $(\bar{x}, f(\bar{x}) + 1) \in \bigcap_{i=1}^m \text{int}(\text{epi}(f_i)) = \bigcap_{i=1}^m \text{ri}(\text{epi}(f_i))$. Furthermore, since $f_i(\bar{x}) < f(\bar{x}) = \bar{\alpha}$ for any $i \notin I(\bar{x})$, there exists a neighborhood U of $\bar{x}$ and $\gamma > 0$ such that $f_i(x) < \alpha$ whenever $(x, \alpha) \in U \times (\bar{\alpha} - \gamma, \bar{\alpha} + \gamma)$. It follows that $(\bar{x}, \bar{\alpha}) \in \text{int}(\text{epi}(f_i))$, and so $N((\bar{x}, \bar{\alpha}); \text{epi}(f_i)) = \{(0, 0)\}$ for such indices i. Thus Theorem 5.3 tells us that

$$N\big((\bar{x}, f(\bar{x})); \text{epi}(f)\big) = \sum_{i=1}^m N\big((\bar{x}, \bar{\alpha}); \text{epi}(f_i)\big) = \sum_{i \in I(\bar{x})} N\big((\bar{x}, f_i(\bar{x})); \text{epi}(f_i)\big).$$

Picking now $v \in \partial f(\bar{x})$, we have by Proposition 6.1(ii) that $(v, -1) \in N((\bar{x}, f(\bar{x})); \text{epi} f)$, which allows us to find $(v_i, -\lambda_i) \in N((\bar{x}, f_i(\bar{x})); \text{epi} f_i)$ for $i \in I(\bar{x})$ such that

$$(v, -1) = \sum_{i \in I(\bar{x})} (v_i, -\lambda_i).$$

This yields $\sum_{i \in I(\bar{x})} \lambda_i = 1$, $\lambda_i \geq 0$, $v = \sum_{i \in I(\bar{x})} v_i$, and $v_i \in \lambda_i \partial f_i(\bar{x})$ by Lemma 8.1(ii). Thus $v = \sum_{i \in I(\bar{x})} \lambda_i u_i$, where $u_i \in \partial f_i(\bar{x})$ and $\sum_{i \in I(\bar{x})} \lambda_i = 1$. This verifies that

$$v \in \text{co}\left(\bigcup_{i \in I(\bar{x})} \partial f_i(\bar{x})\right).$$

25

The opposite inclusion in the maximum rule follows from

$$\partial f_i(\bar{x}) \subset \partial f(\bar{x}) \text{ for all } i \in I(\bar{x}),$$

which in turn follows directly from the definitions. □

9 Optimal Value Function and Another Chain Rule

The main result of this section concerns calculating the subdifferential of extended-real-valued convex functions, which play a remarkable role in variational analysis, optimization, and their numerous applications and are known under the name of *optimal value/marginal functions*. Functions of this class are generally defined by

$$\mu(x) := \inf \{\varphi(x, y) \mid y \in F(x)\}, \tag{9.1}$$

where $\varphi \colon \mathbb{R}^n \times \mathbb{R}^p \to (-\infty, \infty]$ is an extended-real-valued function, and where $F \colon \mathbb{R}^n \rightrightarrows \mathbb{R}^p$ is a set-valued mapping, i.e., $F(x) \subset \mathbb{R}^p$ for every $x \in \mathbb{R}^n$. In what follows we select φ and F in such a way that the resulting function (9.1) is convex and to derive a formula to express its subdifferential via the subdifferential of φ and an appropriate generalized differentiation construction for the set-valued mapping F. The results obtained in the general framework of variational analysis [13, 21] advise us that the most suitable construction for F for these purposes is the so-called *coderivative* of F at $(\bar{x}, \bar{y}) \in \mathrm{gph}(F)$ defined via the normal cone to the graphical set $\mathrm{gph}(F) := \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^p \mid y \in F(x)\}$ by

$$D^* F(\bar{x}, \bar{y})(v) = \{u \in \mathbb{R}^n \mid (u, -v) \in N((\bar{x}, \bar{y}); \mathrm{gph}(F))\}, \quad v \in \mathbb{R}^p. \tag{9.2}$$

It is easy to check that the optimal value function (9.1) is convex provided that φ is convex and the graph of F is convex as well. An example of such F is given by the affine mapping B considered in Lemma 7.1. Note that, as follows directly from Lemma 7.1 and definition (9.2), the coderivative of this mapping is calculated by

$$D^* B(\bar{x}, \bar{y})(v) = A^\top v \text{ with } \bar{y} = B(\bar{x}). \tag{9.3}$$

Now we proceed with calculating the subdifferential of (9.1) via that of φ and the coderivative of F. The results of this type are induced by those in variational analysis [13, 21], where only upper estimate of $\partial \mu(\bar{x})$ were obtained. The convexity setting of this paper and the developed approach allow us to derive an *exact formula* (equality) for calculating $\partial \mu(\bar{x})$ under a mild relative interior condition, which is strictly weaker than the normal qualification condition from [14, Theorem 2.61]; cf. the discussion at the end of Section 5.

Theorem 9.1 *Let $\mu(\cdot)$ be the optimal value function (9.1) generated by a convex-graph mapping $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^p$ and a convex function $\varphi : \mathbb{R}^n \times \mathbb{R}^p \to (-\infty, \infty]$. Suppose that $\mu(x) > -\infty$ for all $x \in \mathbb{R}^n$, fix some $\bar{x} \in \mathrm{dom}(\mu)$, and consider the solution set*

$$S(\bar{x}) := \{\bar{y} \in F(\bar{x}) \mid \mu(\bar{x}) = \varphi(\bar{x}, \bar{y})\}.$$

If $S(\bar{x}) \neq \emptyset$, then for any $\bar{y} \in S(\bar{x})$ we have the equality

$$\partial\mu(\bar{x}) = \bigcup_{(u,v)\in\partial\varphi(\bar{x},\bar{y})} \left[u + D^*F(\bar{x},\bar{y})(v)\right] \tag{9.4}$$

provided the validity of the relative interior qualification condition

$$\mathrm{ri}(\mathrm{dom}(\varphi)) \cap \mathrm{ri}(\mathrm{gph}(F)) \neq \emptyset. \tag{9.5}$$

Proof. Picking any $\bar{y} \in S(\bar{x})$, let us first verify the estimate

$$\bigcup_{(u,v)\in\partial\varphi(\bar{x},\bar{y})} \left[u + D^*F(\bar{x},\bar{y})(v)\right] \subset \partial\mu(\bar{x}). \tag{9.6}$$

To proceed, take w from the set on the left-hand side of (9.6) and find $(u,v) \in \partial\varphi(\bar{x},\bar{y})$ with $w - u \in D^*F(\bar{x},\bar{y})(v)$. It gives us $(w-u,-v) \in N((\bar{x},\bar{y});\mathrm{gph}(F))$ and thus

$$\langle w - u, x - \bar{x}\rangle - \langle v, y - \bar{y}\rangle \leq 0 \ \text{ for all } \ (x,y) \in \mathrm{gph}(F),$$

which shows that whenever $y \in F(x)$ we have

$$\langle w, x - \bar{x}\rangle \leq \langle u, x - \bar{x}\rangle + \langle v, y - \bar{y}\rangle \leq \varphi(x,y) - \varphi(\bar{x},\bar{y}) = \varphi(x,y) - \mu(\bar{x}).$$

This allows us to arrive at the estimate

$$\langle w, x - \bar{x}\rangle \leq \inf_{y\in F(x)} \varphi(x,y) - \mu(\bar{x}) = \mu(x) - \mu(\bar{x})$$

justifying the inclusion $w \in \partial\mu(\bar{x})$ and hence the claimed one in (9.6).

It remains to verify the inclusion "$\subset$" in (9.4). Take $w \in \partial\mu(\bar{x})$, $\bar{y} \in S(\bar{x})$ and get

$$\langle w, x - \bar{x}\rangle \leq \mu(x) - \mu(\bar{x}) = \mu(x) - \varphi(\bar{x},\bar{y}) \leq \varphi(x,y) - \varphi(\bar{x},\bar{y})$$

whenever $y \in F(x)$ and $x \in \mathbb{R}^n$. This implies in turn that for any $(x,y) \in \mathbb{R}^n \times \mathbb{R}^p$ we have

$$\langle w, x - \bar{x}\rangle + \langle 0, y - \bar{y}\rangle \leq \varphi(x,y) + \delta((x,y);\mathrm{gph}(F)) - \left[\varphi(\bar{x},\bar{y}) + \delta((\bar{x},\bar{y});\mathrm{gph}(F))\right].$$

Considering further $f(x,y) := \varphi(x,y) + \delta((x,y);\mathrm{gph}(F))$, deduce from the subdifferential sum rule of Theorem 6.2 under (9.5) that

$$(w,0) \in \partial f(\bar{x},\bar{y}) = \partial\varphi(\bar{x},\bar{y}) + N\big((\bar{x},\bar{y});\mathrm{gph}(F)\big).$$

This shows that $(w,0) = (u_1,v_1)+(u_2,v_2)$ with $(u_1,v_1) \in \partial\varphi(\bar{x},\bar{y})$ and $(u_2,v_2) \in N((\bar{x},\bar{y});\mathrm{gph}(F))$ and thus yields $v_2 = -v_1$. Hence $(u_2,-v_1) \in N((\bar{x},\bar{y});\mathrm{gph}(F))$ meaning by definition that $u_2 \in D^*F(\bar{x},\bar{y})(v_1)$. Therefore we arrive at

$$w = u_1 + u_2 \in u_1 + D^*F(\bar{x},\bar{y})(v_1),$$

which justifies the inclusion "$\subset$" in (9.4) and completes the proof of the theorem. $\qquad\square$

Observe that Theorem 9.1 easily implies the chain rule of Theorem 7.2 by setting $F(x) := \{B(x)\}$ and $\varphi(x, y) := f(y)$ therein. Then we have $\mu(x) = (f \circ B)(x)$,

$$\mathrm{ri}(\mathrm{dom}(\varphi)) = \mathbb{R}^n \times \mathrm{ri}(\mathrm{dom}(f)), \ \mathrm{ri}(\mathrm{gph}(F)) = \mathrm{gph}(B),$$

Thus the relative interiority assumption of Theorem 7.2 yields the validity of the qualification condition (9.5) is satisfied, and we arrive at the chain rule (7.2) directly from (9.4) and the coderivative expression in (9.3).

We now derive from Theorem 9.1 and the intersection rule of Theorem 5.3 a new subdifferential chain rule concerning compositions of convex functions with particular structures. We say that $g : \mathbb{R}^p \to (-\infty, \infty]$ is *nondecreasing componentwise* if

$$\big[x_i \le y_i \ \text{ for all } \ i = 1, \ldots, p\big] \implies \big[g(x_1, \ldots, x_p) \le g(y_1, \ldots, y_p)\big].$$

Theorem 9.2 *Define* $h : \mathbb{R}^n \to \mathbb{R}^p$ *by* $h(x) := (f_1(x), \ldots, f_p(x))$, *where* $f_i : \mathbb{R}^n \to \mathbb{R}$ *for* $i = 1, \ldots, p$ *are convex functions. Suppose that* $g : \mathbb{R}^p \to (-\infty, \infty]$ *is convex and nondecreasing componentwise. Then the composition* $g \circ h : \mathbb{R}^n \to (-\infty, \infty]$ *is a convex function, and we have the subdifferential chain rule*

$$\partial(g \circ h)(\bar{x}) = \Big\{ \sum_{i=1}^{p} \gamma_i v_i \ \Big| \ (\gamma_1, \ldots, \gamma_p) \in \partial g(\bar{y}), \ v_i \in \partial f_i(\bar{x}), \ i = 1, \ldots, p \Big\} \qquad (9.7)$$

with $\bar{x} \in \mathbb{R}^n$ *and* $\bar{y} := h(\bar{x}) \in \mathrm{dom}(g)$ *under the condition that there exist* $\bar{u} \in \mathbb{R}^n$ *and* $\bar{\lambda}_i > f_i(\bar{u})$ *for all* $i = 1, \ldots, p$ *satisfying*

$$(\bar{\lambda}_1, \ldots, \bar{\lambda}_p) \in \mathrm{ri}(\mathrm{dom}(g)).$$

Proof. Let $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^p$ be a set-valued mapping defined by

$$F(x) := [f_1(x), \infty) \times [f_2(x), \infty) \times \ldots \times [f_p(x), \infty).$$

Then the graph of F is represented by

$$\mathrm{gph}(F) = \big\{(x, t_1, \ldots, t_p) \in \mathbb{R}^n \times \mathbb{R}^p \big| \ t_i \ge f_i(x)\big\}.$$

Consider further the convex sets

$$\Omega_i := \{(x, \lambda_1, \ldots, \lambda_p) \mid \lambda_i \ge f_i(x)\}$$

and observe that $\mathrm{gph}(F) = \bigcap_{i=1}^{p} \Omega_i$. Since all the functions f_i are convex, the set $\mathrm{gph}(F)$ is convex as well. Define $\varphi : \mathbb{R}^n \times \mathbb{R}^p \to (-\infty, \infty]$ by $\varphi(x, y) := g(y)$ and observe, since g is increasing componentwise, that

$$\inf \big\{\varphi(x, y) \ \big| \ y \in F(x)\big\} = g\big(f_1(x), \ldots, f_p(x)\big) = (g \circ h)(x),$$

which ensures the convexity of the composition $g \circ h$; see [14, Proposition 1.54]. It follows from Proposition 3.5 and Lemma 5.2 that

$$\mathrm{ri}(\mathrm{gph}(F)) = \{(x, \lambda_1, \ldots, \lambda_p) \mid \lambda_i > f_i(x) \text{ for all } \ i = 1, \ldots, p\}.$$

The assumptions made in the theorem guarantee that $\mathrm{ri}(\mathrm{gph}(F)) \cap \mathrm{ri}(\mathrm{dom}(\varphi)) \neq \emptyset$. Moreover, the structure of each set Ω_i gives us

$$[(v, -\gamma_1, \ldots, -\gamma_p) \in N((\bar{x}, f_1(\bar{x}), \ldots, f_p(\bar{x})); \Omega_i)] \iff [(v, -\gamma_i) \in N((\bar{x}, f_i(\bar{x})); \mathrm{epi}(f_i)), \gamma_j = 0 \text{ if } j \neq i].$$

Using the coderivative definition and applying the intersection rule of Theorem 5.3 and also Lemma 8.1, we get

$$v \in D^* F(\bar{x}, f_1(\bar{x}), \ldots, f_p(\bar{x}))(\gamma_1, \ldots, \gamma_p) \iff (v, -\gamma_1, \ldots, -\gamma_p) \in N((\bar{x}, f_1(\bar{x}), \ldots, f_p(\bar{x})); \mathrm{gph}(F)))$$

$$\iff (v, -\gamma_1, \ldots, -\gamma_p) \in N((\bar{x}, f_1(\bar{x}), \ldots, f_p(\bar{x})); \bigcap_{i=1}^{p} \Omega_i)$$

$$\iff (v, -\gamma_1, \ldots, -\gamma_p) \in \sum_{i=1}^{p} N((\bar{x}, f_1(\bar{x}), \ldots, f_p(\bar{x})); \Omega_i)$$

$$\iff v = \sum_{i=1}^{p} v_i \text{ with } (v_i, -\gamma_i) \in N((\bar{x}, f_i(\bar{x})); \mathrm{epi} f_i)$$

$$\iff v = \sum_{i=1}^{p} v_i \text{ with } v_i \in \gamma_i \partial f_i(\bar{x})$$

$$\iff v \in \sum_{i=1}^{p} \gamma_i \partial f_i(\bar{x}).$$

It follows from Theorem 9.1 that $v \in \partial(g \circ h)(\bar{x})$ if and only if there exists a collection $(\gamma_1, \ldots, \gamma_p) \in \partial g(\bar{y})$ such that $v \in D^* F(\bar{x}, f_1(\bar{x}), \ldots, f_p(\bar{x}))(\gamma_1, \ldots, \gamma_p)$. This allows us to deduce the chain rule (9.7) from the equivalences above. $\qquad\square$

10 Normals to Preimages of Sets via Set-Valued Mappings

In this section we calculate the normal cone to convex sets of a special structure that frequently appear in variational analysis and optimization. These sets are constructed as follows. Given a set $\Theta \subset \mathbb{R}^p$ and a set-valued mapping $F \colon \mathbb{R}^n \rightrightarrows \mathbb{R}^p$ the *preimage* or *inverse image* of the set Θ under the mapping F is defined by

$$F^{-1}(\Theta) := \{x \in \mathbb{R}^n \mid F(x) \cap \Theta \neq \emptyset\}. \tag{10.1}$$

Our goal here is to calculate the normal cone to the preimage set (10.1) via the normal cone to Θ and the coderivative of F. This is done in the following theorem, which is yet another consequence of the intersection rule from Theorem 5.3.

Theorem 10.1 *Let $F \colon \mathbb{R}^n \rightrightarrows \mathbb{R}^p$ be a set-valued mapping with convex graph, and let Θ be a convex subset of $\mathbb{R}^p$. Suppose that there exists $(a, b) \in \mathbb{R}^n \times \mathbb{R}^p$ satisfying*

$$(a, b) \in \mathrm{ri}(\mathrm{gph}(F)) \text{ and } b \in \mathrm{ri}(\Theta).$$

Then for any $\bar{x} \in F^{-1}(\Theta)$ and $\bar{y} \in F(\bar{x}) \cap \Theta$ we have the representation

$$N(\bar{x}; F^{-1}(\Theta)) = D^* F(\bar{x}, \bar{y})(N(\bar{y}; \Theta)). \tag{10.2}$$

Proof. It is not hard to show that $F^{-1}(\Theta)$ is a convex set. Picking any $u \in N(\bar{x}; F^{-1}(\Theta))$ gives us by definition that

$$\langle u, x - \bar{x} \rangle \leq 0 \text{ whenever } x \in F^{-1}(\Theta), \text{ i.e., } F(x) \cap \Theta \neq \emptyset.$$

Consider the two convex subsets of $\mathbb{R}^{n+p}$ defined by

$$\Omega_1 := \text{gph}(F) \text{ and } \Omega_2 := \mathbb{R}^n \times \Theta$$

for which we have $(u, 0) \in N((\bar{x}, \bar{y}); \Omega_1 \cap \Omega_2)$. Applying now Theorem 5.3 tells us that

$$(u, 0) \in N((\bar{x}, \bar{y}); \Omega_1) + N((\bar{x}, \bar{y}); \Omega_2) = N((\bar{x}, \bar{y}); \text{gph}F) + [\{0\} \times N(\bar{y}; \Theta)],$$

and thus we get the representation

$$(u, 0) = (u, -v) + (0, v) \text{ with } (u, -v) \in N((\bar{x}, \bar{y}); \text{gph}F) \text{ and } v \in N(\bar{y}; \Theta)$$

from which it follows immediately that

$$u \in D^*F(\bar{x}, \bar{y})(v) \text{ and } v \in N(\bar{y}; \Theta).$$

This versifies the inclusion "$\subset$" in (10.2). The opposite inclusion is trivial. $\qquad\square$

11 Coderivative Calculus

We see from above that the coderivative notion (9.2) is instrumental to deal with set-valued mappings. Although this notion was not properly developed in basic convex analysis, the importance of it has been fully revealed in general variational analysis and its applications; see, e.g., [4, 13, 21] and the references therein, where the reader can find, in particular, various results on coderivative calculus. Most of these results were obtained in the inclusion form under the corresponding normal qualification conditions generated by (5.8). We present below some calculus rules for coderivatives of convex-graph mappings, which are derived from the intersection rule of Theorem 5.3 and hold as *equalities* in a bit different form under weaker *relative interior qualification conditions*.

Recall that the domain of a set-valued mapping $F \colon \mathbb{R}^n \rightrightarrows \mathbb{R}^p$ is defined by

$$\text{dom}(F) := \{x \in \mathbb{R}^m \mid F(x) \neq \emptyset\}.$$

Given two set-valued mappings $F_1, F_2 \colon \mathbb{R}^n \rightrightarrows \mathbb{R}^p$, their sum is defined by

$$(F_1 + F_2)(x) = F_1(x) + F_2(x) := \{y_1 + y_2 \mid y_1 \in F_1(x), \ y_2 \in F_2(x)\}.$$

It is easy to see that $\text{dom}(F_1 + F_2) = \text{dom}(F_1) \cap \text{dom}(F_2)$ and that the graph of $F_1 + F_2$ is convex provided that both F_1, F_2 have this property.

Our first calculus result concerns representing the coderivative of the sum $F_1 + F_2$ at the given point $(\bar{x}, \bar{y}) \in \text{gph}(F_1 + F_2)$. To formulate it, consider the nonempty set

$$S(\bar{x}, \bar{y}) := \{(\bar{y}_1, \bar{y}_2) \in \mathbb{R}^p \times \mathbb{R}^p \mid \bar{y} = \bar{y}_1 + \bar{y}_2, \ \bar{y}_i \in F_i(\bar{x}) \text{ for } i = 1, 2\}.$$

Theorem 11.1 *Let $F_1, F_2 \colon \mathbb{R}^n \rightrightarrows \mathbb{R}^p$ be set-valued mappings of convex graphs, and let the relative interior qualification condition*

$$\mathrm{ri}(\mathrm{gph}(F_1)) \cap \mathrm{ri}(\mathrm{gph}(F_2)) \neq \emptyset \qquad (11.1)$$

hold. Then we have the coderivative sum rule

$$D^*(F_1 + F_2)(\bar{x}, \bar{y})(v) = \bigcap_{(\bar{y}_1, \bar{y}_2) \in S(\bar{x}, \bar{y})} [D^* F_1(\bar{x}, \bar{y}_1)(v) + D^* F_2(\bar{x}, \bar{y}_2)(v)] \qquad (11.2)$$

for all $(\bar{x}, \bar{y}) \in \mathrm{gph}(F_1 + F_2)$ and $v \in \mathbb{R}^p$.

Proof. Fix any $u \in D^*(F_1 + F_2)(\bar{x}, \bar{y})(v)$ and $(\bar{y}_1, \bar{y}_2) \in S(\bar{x}, \bar{y})$ for which we have the inclusion $(u, -v) \in N((\bar{x}, \bar{y}); \mathrm{gph}(F_1 + F_2))$. Consider the convex sets

$$\Omega_1 := \{(x, y_1, y_2) \in \mathbb{R}^n \times \mathbb{R}^p \times \mathbb{R}^p \mid y_1 \in F_1(x)\}, \quad \Omega_2 := \{(x, y_1, y_2) \in \mathbb{R}^n \times \mathbb{R}^p \times \mathbb{R}^p \mid y_2 \in F_2(x)\}$$

and deduce from the normal cone definition that

$$(u, -v, -v) \in N((\bar{x}, \bar{y}_1, \bar{y}_2); \Omega_1 \cap \Omega_2).$$

It is easy to observe the relative interior representations

$$\mathrm{ri}(\Omega_1) = \{(x, y_1, y_2) \in \mathbb{R}^n \times \mathbb{R}^p \times \mathbb{R}^p \mid (x, y_1) \in \mathrm{ri}(\mathrm{gph}(F_1))\},$$
$$\mathrm{ri}(\Omega_2) = \{(x, y_1, y_2) \in \mathbb{R}^n \times \mathbb{R}^p \times \mathbb{R}^p \mid (x, y_2) \in \mathrm{ri}(\mathrm{gph}(F_2))\},$$

which show that condition (11.1) yields $\mathrm{ri}(\Omega_1) \cap \mathrm{ri}(\Omega_2) \neq \emptyset$. It tells us by Theorem 5.3 that

$$(u, -v, -v) \in N((\bar{x}, \bar{y}_1, \bar{y}_2); \Omega_1) + N((\bar{x}, \bar{y}_1, \bar{y}_2); \Omega_2),$$

and thus we arrive at the representation

$$(u, -v, -v) = (u_1, -v, 0) + (u_2, 0, -v) \quad \text{with} \quad (u_i, -v) \in N((\bar{x}, \bar{y}_i); \mathrm{gph}(F_i)), \ i = 1, 2.$$

This verifies therefore the relationship

$$u = u_1 + u_2 \in D^* F_1(\bar{x}, \bar{y}_1)(v) + D^* F_2(\bar{x}, \bar{y}_2)(v),$$

which justifies the inclusion "$\subset$" in (11.2). The opposite inclusion is obvious. $\qquad \square$

Next we define the composition of two mappings $F \colon \mathbb{R}^n \rightrightarrows \mathbb{R}^p$ and $G \colon \mathbb{R}^p \rightrightarrows \mathbb{R}^q$ by

$$(G \circ F)(x) = \bigcup_{y \in F(x)} G(y) := \{z \in G(y) \mid y \in F(x)), \quad x \in \mathbb{R}^n,$$

and observe that $G \circ F$ is convex-graph provided that both F and G have this property. Given $\bar{z} \in (G \circ F)(\bar{x})$, we consider the set

$$M(\bar{x}, \bar{z}) := F(\bar{x}) \cap G^{-1}(\bar{z}).$$

Theorem 11.2 *Let $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^p$ and $G : \mathbb{R}^p \rightrightarrows \mathbb{R}^q$ be set-valued mappings of convex graphs for which there exist vectors $(x, y, z) \in \mathbb{R}^n \times \mathbb{R}^p \times \mathbb{R}^q$ satisfying*

$$(x, y) \in \mathrm{ri}(\mathrm{gph}(F)) \text{ and } (y, z) \in \mathrm{ri}(\mathrm{gph}(G)). \tag{11.3}$$

Then for any $(\bar{x}, \bar{z}) \in \mathrm{gph}(G \circ F)$ and $w \in \mathbb{R}^q$ we have the coderivative chain rule

$$D^*(G \circ F)(\bar{x}, \bar{z})(w) = \bigcap_{\bar{y} \in M(\bar{x}, \bar{z})} D^* F(\bar{x}, \bar{y}) \circ D^* G(\bar{y}, \bar{z})(w). \tag{11.4}$$

Proof. Picking $u \in D^*(G \circ F)(\bar{x}, \bar{z})(w)$ and $\bar{y} \in M(\bar{x}, \bar{z})$ gives us the inclusion $(u, -w) \in N((\bar{x}, \bar{z}); \mathrm{gph}(G \circ F))$, which means that

$$\langle u, x - \bar{x} \rangle - \langle w, z - \bar{z} \rangle \leq 0 \text{ for all } (x, z) \in \mathrm{gph}(G \circ F).$$

Form now the two convex subsets of $\mathbb{R}^n \times \mathbb{R}^p \times \mathbb{R}^q$ by

$$\Omega_1 := \mathrm{gph}(F) \times \mathbb{R}^q \text{ and } \Omega_2 := \mathbb{R}^n \times \mathrm{gph}(G)$$

We can easily deduce from the definitions that

$$(u, 0, -w) \in N((\bar{x}, \bar{y}, \bar{z}); \Omega_1 \cap \Omega_2)$$

and that the qualification condition (11.3) ensures the validity of the one $\mathrm{ri}(\Omega_1) \cap \mathrm{ri}(\Omega_2) \neq \emptyset$ in Theorem 5.3. Applying then the intersection rule to the above sets tells us that

$$(u, 0, -w) \in N((\bar{x}, \bar{y}, \bar{z}); \Omega_1 \cap \Omega_2) = N((\bar{x}, \bar{y}, \bar{z}); \Omega_1) + N((\bar{x}, \bar{y}, \bar{z}); \Omega_2),$$

and thus there is a vector $v \in \mathbb{R}^p$ such that we have the representation

$$(u, 0, -w) = (u, -v, 0) + (0, v, -w) \text{ with } (u, -v) \in N((\bar{x}, \bar{y}); \mathrm{gph}(F)), \ (v, -w) \in N((\bar{y}, \bar{z}); \mathrm{gph}(G)).$$

This shows by the coderivative definition (9.2) that

$$u \in D^* F(\bar{x}, \bar{y})(v) \text{ and } v \in D^* G(\bar{y}, \bar{z})(w),$$

and so we justify the inclusion "$\subset$" in (11.4). The opposite inclusion is easy to verify. $\square$

Our final result in this section provides an exact formula for calculating the coderivative of intersections of set-valued mappings $F_1, F_2 : \mathbb{R}^n \rightrightarrows \mathbb{R}^p$ defined by

$$(F_1 \cap F_2)(x) := F_1(x) \cap F_2(x), \quad x \in \mathbb{R}^n,$$

which is also deduced from the basic intersection rule for the normal cone in Theorem 5.3.

Proposition 11.3 *Let $F_1, F_2 : \mathbb{R}^n \rightrightarrows \mathbb{R}^p$ be of convex graphs, and let*

$$\mathrm{ri}(\mathrm{gph}(F_1)) \cap \mathrm{ri}(\mathrm{gph}(F_2)) \neq \emptyset.$$

Then for any $\bar{y} \in (F_1 \cap F_2)(\bar{x})$ and $v \in \mathbb{R}^p$ we have

$$D^*(F_1 \cap F_2)(v) = \bigcup_{v_1 + v_2 = v} \left[D^* F_1(\bar{x}, \bar{y})(v_1) + D^* F_2(\bar{x}, \bar{y})(v_2) \right]. \tag{11.5}$$

Proof. It follows from the definition that $\mathrm{gph}(F_1 \cap F_2) = \mathrm{gph}(F_1) \cap \mathrm{gph}(F_2)$. Pick any vector $u \in D^*(F_1 \cap F_2)(v)$ and get by Theorem 5.3 that

$$(u, -v) \in N((\bar{x}, \bar{y}); \mathrm{gph}(F_1 \cap F_2)) = N((\bar{x}, \bar{y}); \mathrm{gph}(F_1)) + N((\bar{x}, \bar{y}); \mathrm{gph}(F_2)).$$

This allows us to represent the pair (u, v) in the form

$$(u, -v) = (u_1, -v_1) + (u_2, -v_2),$$

where $(u_1, -v_1) \in N((\bar{x}, \bar{y}); \mathrm{gph}(F_1))$ and $(u_2, -v_2) \in N((\bar{x}, \bar{y}); \mathrm{gph}(F_2))$. Therefore we have

$$u = u_1 + u_2 \in D^*F_1(\bar{x}, \bar{y})(v_1) + D^*F_2(\bar{x}, \bar{y})(v_2) \ \text{ with } \ v = v_1 + v_2,$$

verifying the inclusion "$\subset$" in (11.5). The opposite inclusion comes from the definition. $\square$

12 Solution Maps for Parameterized Generalized Equations

Here we present a rather simple application of coderivative calculus to calculating the coderivative of set-valued mappings given in the structural form

$$S(x) = \{y \in \mathbb{R}^p \mid 0 \in F(x, y) + G(x, y)\}, \quad x \in \mathbb{R}^n, \tag{12.1}$$

where $F, G : \mathbb{R}^n \times \mathbb{R}^p \rightrightarrows \mathbb{R}^q$ are set-valued mappings. Mappings of this type can be treated as solutions maps to the so-called *generalized equations*

$$0 \in F(x, y) + G(x, y), \quad x \in \mathbb{R}^n, \ y \in \mathbb{R}^p,$$

with respect to the decision variable y under parameterization/perturbation by x. This terminology and first developments go back to Robinson [18], who considered the case where $G(y) = N(y; \Omega)$ is the normal cone mapping associated with a convex set Ω and where $F(x, y)$ is single-valued. It has been recognized that the generalized equation formalism, including its extended form, is a convenient model to investigate various aspects of optimization, equilibrium, stability, etc. In particular, the coderivative of the solution map (12.1) plays an important role in such studies; see, e.g., [13] and the references therein.

To proceed with calculating the coderivative of the solution map (12.1) in the case of convex-graph set-valued mappings, we first observe the following fact of its own interest.

Proposition 12.1 *Let $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^p$ be an arbitrary set-valued mapping with convex graph. Given $\bar{x} \in \mathrm{dom}(F)$, we have the relationships*

$$N(\bar{x}; \mathrm{dom}(F)) = D^*F(\bar{x}, \bar{y})(0) \ \text{ for every } \ \bar{y} \in F(\bar{x}). \tag{12.2}$$

Proof. Picking any $v \in N(\bar{x}; \mathrm{dom}(F))$ and $\bar{y} \in F(\bar{x})$ gives us

$$\langle v, x - \bar{x} \rangle \leq 0 \text{ for all } \ x \in \mathrm{dom}(F),$$

33

which immediately implies the inequality

$$\langle v, x - \bar{x}\rangle + \langle 0, y - \bar{y}\rangle \leq 0, \quad y \in F(x).$$

This yields in turn $(v, 0) \in N((\bar{x}, \bar{y}); \mathrm{gph})(F))$ and so $v \in D^*F(\bar{x}, \bar{y})(0)$ thus verifying the inclusion "⊂" in (12.2). The opposite inclusion in (12.2) is straightforward. □

Theorem 12.2 *Let F and G in (12.1) be convex-graph, and let $(\bar{x}, \bar{y}) \in \mathrm{gph}S$. Impose the qualification condition*

$$\mathrm{ri}(\mathrm{gph}(F)) \cap \mathrm{ri}(-\mathrm{gph}(G)) \neq \emptyset. \tag{12.3}$$

Then for every $\bar{z} \in F(\bar{x}, \bar{y}) \cap [-G(\bar{x}, \bar{y})]$ we have

$$D^*S(\bar{x}, \bar{y})(v) = \bigcup_{w \in \mathbb{R}^q} \{u \in \mathbb{R}^n \mid (u, -v) \in D^*F((\bar{x}, \bar{y}), \bar{z})(w) + D^*G((\bar{x}, \bar{y}), -\bar{z})(w)\}. \tag{12.4}$$

Proof. It is easy to see that the solution map S is convex-graph under this property for F and G. Furthermore, we get

$$\begin{aligned}
\mathrm{gph}(S) &= \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^p \mid 0 \in F(x, y) + G(x, y)\} \\
&= \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^p \mid F(x, y) \cap [-G(x, y)] \neq \emptyset\} = \mathrm{dom}(H),
\end{aligned}$$

where $H(x, y) := F(x, y) \cap [-G(x, y)]$. Take further any $u \in D^*S(\bar{x}, \bar{y})(v)$ and deduce from (9.2) and Proposition 12.1 that

$$(u, -v) \in N((\bar{x}, \bar{y}); \mathrm{gph}(S)) = N((\bar{x}, \bar{y}); \mathrm{dom}(H)) = D^*H((\bar{x}, \bar{y}, \bar{z}))(0)$$

for every $\bar{z} \in H(\bar{x}, \bar{y}) = F(\bar{x}, \bar{y}) \cap [-G(\bar{x}, \bar{y})]$. Then the coderivative intersection rule of Proposition 11.3 tells us that under the validity of (12.3) that

$$\begin{aligned}
(u, -v) \in D^*H((\bar{x}, \bar{y}, \bar{z}))(0) &= \bigcup_{w \in \mathbb{R}^q} [D^*F((\bar{x}, \bar{y}), \bar{z})(w) + D^*(-G)((\bar{x}, \bar{y}), \bar{z})(-w)] \\
&= \bigcup_{w \in \mathbb{R}^q} [D^*F((\bar{x}, \bar{y}), \bar{z})(w) + D^*(G)((\bar{x}, \bar{y}), -\bar{z})(w)],
\end{aligned}$$

which yields (12.4) and thus completes the proof of the theorem. □

References

[1] H. H. Bauschke and P. L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, Springer, New York, 2011.

[2] D. P. Bertsekas, A. Nedić and A. E. Ozdaglar, *Convex Analysis and Optimization*, Athena Scientific, Belmont, MA, 2003.

[3] J. M. Borwein and A. S. Lewis, *Convex Analysis and Nonlinear Optimization*, Springer, New York, 2000.

[4] J. M. Borwein and Q. J. Zhu, *Techniques of Variational Analysis*, Springer, New York, 2005.

[5] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, New York, 2004.

[6] W. Fenchel, *Convex Cones, Sets and Functions*, Lecture Notes, Princeton University, Princeton, NJ, 1951.

[7] F. Giannessi, *Constrained Optimization and Image Space Analysis, I: Separation of Sets and Optimality Conditions*, Springer, Berlin, 2005.

[8] J.-B. Hiriart-Urruty and C. Lemaréchal, *Fundamentals of Convex Analysis*, Springer, Berlin, 2001.

[9] J.-B. Hiriart-Urruty and C. Lemaréchal, *Convex Analysis and Minimization Algorithms, I: Fundamentals*, Springer, Berlin, 1993.

[10] A. G. Kusraev and S. S. Kutateladze, *Subdifferentials: Theory and Applications*, Kluwer, Dordrecht, The Netherlands, 1995.

[11] G. G. Magaril-Il'yaev and V. M. Tikhomirov, *Convex Analysis: Theory and Applications*, American Mathematical Society, Providence, RI, 2003.

[12] H. Minkowski, *Geometrie der Zahlen*, Teubner, Leipzig, 1910.

[13] B. S. Mordukhovich, *Variational Analysis and Generalized Differentiation, I: Basic Theory, II: Applications*, Springer, Berlin, 2006.

[14] B. S. Mordukhovich and N. M. Nam, *An Easy Path to Convex Analysis and Applications*, Morgan & Claypool Publishers, San Rafael, CA, 2014.

[15] J. J. Moreau, Propriétés des applications prox, *C. R. Acad. Sci. Paris* **256** (1963), 1069–1071.

[16] D. Pallaschke and S. Rolewicz, *Foundation of Mathematical Optimization: Convex Analysis without Linearity*, Kluwer, Dordrecht, The Netherlands, 1998.

[17] B. T. Polyak, *Introduction to Optimization*, Optimization Software, New York, 1987.

[18] S. M. Robinson, Generalized equations and their solutions, I: Basic theory, *Math. Program. Stud.* **10** (1979), 128–141.

[19] R. T. Rockafellar, *Convex Functions and Dual Extremum Problems*, Ph.D. disseration, Harvard University, Cambridge, MA, 1963.

[20] R. T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, NJ, 1970.

[21] R. T. Rockafellar and R. J-B. Wets. *Variational Analysis*, Springer, Berlin, 1998.

[22] C. Zălinescu, *Convex Analysis in General Vector Spaces,* World Scientific, Singapore, 2002.